The Design and Analysis of Post-Epidemic Foot-and-Mouth Disease Surveillance Programmes.

Ian Handel

PhD

THE UNIVERSITY OF EDINBURGH

2010

Declaration of Authorship

I, Ian Graham Handel, declare that this thesis titled, 'The design and analysis of post-epidemic foot-and-mouth disease surveillance programmes' and the work presented in it are my own. I confirm that:

- No part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date: 6/10/2010

Abstract

Post-epidemic surveillance is a necessary component in the process by which areas having experienced epidemic or endemic animal disease gain a 'disease-free' status. Whilst it is useful to rapidly and surely regain this status, surveillance uses expensive resources and imposes a time delay on regaining disease-freedom. It is beneficial to design and analyse surveillance to maximise efficiency and use data effectively. This thesis addresses issues of surveillance design and analysis in the context of demonstration of disease freedom after foot-and-mouth disease (FMD) epidemics or outbreaks. International regulations require the process of demonstrating disease freedom to include a serological survey of livestock after the epidemic has apparently ended. Currently, animal holdings are sampled from randomly to achieve a defined probability of detecting disease, if present in the region. Using a risk model applied to demographic and epidemic data from the Devon UK 2001 FMD outbreak, I estimate the efficiency gains of risk based sampling compared to random sampling. With this technique farms at a high risk of harbouring undiscovered infection are selectively targeted for sampling. This approach is robust to model errors and reduces the number of farms to be sampled from 1083 to 225 to achieve a 95% level of confidence that the area is disease free. Additionally, using the risk model to order the sampling of farms will reduce delays to declaration of disease-freedom by approximately 11 days on average. The Thrace region of Turkey is a buffer zone between the rest of Europe, which is generally FMD free, and Anatolian Turkey, which has endemic FMD. A EU/FAO funded project vaccinates and serologically tests Thrace's livestock. This provides a data set for the exploration of survey analysis techniques. The testing is for the purpose of demonstrating disease freedom so Turkey may join the European Union. Using a stochastic model simulating different disease scenarios and diagnostic test performances I evaluate the current sampling strategy. It generally satisfies the sensitivity requirement to detect disease in the area. However the initial specificity is low, resulting in a high rate of false positive classifications of villages, requiring the use of confirmatory tests. I describe the results from three surveys in 2005-2006 and adjust for an apparent shift in test result datum over the period. The results are then used to parameterise a multilevel, mixture model of FMD in Thrace. Estimates from this model suggest a high village-level prevalence of disease with high variability of within-village prevalence. The temporal changes are compatible with the reported FMD outbreaks though the absolute estimates are much higher than expected from outbreak data. I suggest that there may be relatively few truly exposed villages with the large number of apparently exposed villages the consequence of imperfect diagnostic test specificity in the vaccinated population. Post-epidemic surveillance relies on serological and clinical surveillance. I develop a simulation model based on a within-group stochastic model and models of stockholder observation and diagnostic test response to estimate the time varying results of clinical surveillance and the additional benefits of serological surveillance. This quantifies current belief that serological surveillance should be focused on more extensive production systems. The additional benefits of serological surveillance are greatest in poorly observed, sparse outbreaks allowing, in the simulated scenarios, disease freedom to be declared in animal groups 15-33 days sooner than by clinical observation alone.

Acknowledgements

I would like to thank everyone that has helped, advised and supported me with this PhD thesis.

I thank supervisors and project PI's Mark Woolhouse and Mark Bronsvoort for having confidence that a veterinary surgeon from general practice can transform into a quantitative epidemiologist. I am grateful for their support and encouragement. I would also like to thank my third supervisor, John Forbes, for his valuable input, constructive criticism and support.

My study was carried out part-time during a fellowship funded by DEFRA and the Scottish Funding Council under grant VT 0101. Across the UK VTRI has been a successful initiative and I'm proud to be a small part of it. The collaborations within our VTRI project and with other projects have been very valuable both as learning and research processes. Being a member of the Edinburgh/Glasgow VTRI project has also provided a community of colleagues and friends for support. They include: Lisa Boden, Eric Fèvre, Emily Courcier, Fransceca Scholmacchi, Margo Chase-Topping, Giles Innocent, Jane Heller, Tim Parkin, Matthew Denwood, Stuart Reid and Helen Ternent.

As a VTRI fellow I had the benefit of a steering panel out-with my PhD project. This panel oversaw both the science of my research and the wider educational and policy relevance of my activities. I am grateful for the time and effort spent by the further members of this panel which included David Black, David Paton and Matt Keeling. For the duration of this PhD I have been a member of the Epigroup epidemiology research group. I would like to thank all members past and present for their friendship and support. I am particularly grateful to past members to Shaun Abeysinghe and Louise Matthews for their patient support and advice.

One of the advantages of being a PhD student is that others are exceptionally generous with their time. I would like to thank everyone that has listened to my questions, refined them and then helped answer them. These people include many of the above and additionally Paul Bessell, Peter Diggle, Peter Cripps, Iain McKendrick, Nick Savill and Satya Parida.

Chapter 6 involved development of a model of on-farm disease detection informed by literature review and interviews with stake-holders. This inputs were gathered as part of a VTRI summer studentship by an undergraduate veterinary student, Cecilia Kung. Cecilia's meticulous note-keeping and enthusiasm for the project provided great support at a time when a PhD thesis seemed to be a distant dream. I would also like to thank the clinicians, stock-keepers and veterinary surgeons that consented to interviews for this study. Kim Willoughby and Ruth Zadoks from the Moredun Institute generously gave me access to their literature review concerning the transmission and persistence of foot-and-mouth disease.

The studies of foot-and-mouth disease sero-surveillance design and analysis in Thrace, Turkey relied heavily on co-operation and data from both the European foot-and-mouth disease research Group (EUFMD) and epidemiologist colleagues in Turkey. I would like to thank all the members of the EUFMD including Keith Sumption, Kris DeClerq and Tom Murray together with Naci Bulut from The ŞAP Institute in Turkey for their kind hospitality, help and support.

Outside of Edinburgh I would like to acknowledge the early career support of Barbara Weaver. An incredible and determined veterinary anaesthetist who had to suffer my first meanderings in mathematical modelling over twenty years ago when I was an undergraduate. Barbara's enthusiasm for rigorous, quantitative research was probably one of the main drivers to my return to this area decades later.

The mates and colleagues blur a lot with the main partners in crime outside of work. Mark Bronsvoort has done a hero's job of being a PhD supervisor whilst suffering hills, blizzards, cold and rain on Scottish mountains and bike trails. Likewise Neil Hawkins, an inspirational statistician and health economist who always found time to talk about work when we really should have been worrying about how to get back off the hills by closing time. Mihal Cotter who shared tea breaks and beans on toast while I wrote up. Massive thanks also to David Spence, mate from way back, who supported my apparently insane move from business to academia with enthusiasm and encouragement.

Last I'd like to take a rare opportunity to thank my family: My parents and sister have been great all the way with every bit of school, work, play and life. I couldn't ask for more. Rachel, my partner and Joshua and Cameron our boys – there's nothing I can write here that does you justice. I bet you're just glad it's written.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

This thesis considers elements of the design and analysis of post-epidemic surveillance in the context of foot-and-mouth disease. This chapter provides a background on general issues of disease surveillance and the specific issues pertaining to post-epidemic foot-and-mouth disease surveillance.

Chapter 2 considers a risk model based approach to post-epidemic serological surveillance for foot-and-mouth disease. In this I propose that the aims of post-epidemic surveillance may be more economically and more quickly met by applying risk based rather than random selection based criteria to the surveillance of a previously diseased region.

The Thrace region of Turkey is a buffer between the, foot-and-mouth disease endemic, Anatolian region of Turkey and, foot-and-mouth disease-free, Europe. There is an ongoing vaccination and serological testing programme in Thrace to control foot-and-mouth disease incursions and to monitor foot-and-mouth disease control. In chapter 3, I use a simulation model to evaluate the probable efficacy of the ongoing testing programme. Data are available from three region-wide serological surveys. In chapter 4, I explore these data and apply a calibration to the diagnostic test results and in chapter 5, I use a Bayesian, multilevel, mixture model to make inferences about the epidemiology of foot-and-mouth disease in Thrace. I also discuss the limitations of using data from disease-freedom surveillance for more general inference.

Alongside serological surveillance there will always be implicit clinical surveillance of livestock after a disease outbreak or epidemic. In chapter 6, I combine individually based stochastic models of within herd epidemiology with simulation models of clinical and serological surveillance to examine the relative performance of these systems.

Chapter 7, the general discussion, collates the major findings of the thesis and considers their limitations, implications and consequences as well as opportunities for further study.

1.2 Foot-and-mouth disease

1.2.1 History

Foot-and-mouth disease was first described in 1546 by a monk, Hieronymous Fracastorius, in his observations of an epidemic in the cattle near Verona, Italy [Mahy, 2005]. The causative agent, foot-and-mouth disease virus was discovered in the late nineteenth century by Friedrich Loeffler and Paul Frosch [Mahy, 2005]. They observed that fluid from infected animals remained infectious when passed through sub-bacteria sized filters. This was the first recorded description of a viral disease of animals. Foot-and-mouth disease occurs throughout much of the world, including parts of Europe, Africa, Asia, and South America. In the UK it was first recorded in 1839 where it remained endemic until the 1880s. After a period of absence, foot-and-mouth disease returned to the UK in the year 1900 with a total of 21 outbreaks followed by a further 12 outbreaks in 1901 and a single case in 1902. There was then a period of disease-freedom until 1908, with a major epidemic in 1922-24. Foot-and-mouth disease remained present in the UK until 1962. During this period, investigation of transmission of the foot-and-mouth disease virus informed a number of control policies including restriction of meat imports from endemic countries and regulations to require the boiling of food waste fed as swill to pigs. These strategies, in combination with control measures in neighbouring countries, eliminated foot-and-mouth disease from the UK by 1962 [DEFRA, 2009]. Since then there have been two major UK epidemics; one in 1967-68, thought have arisen from the feeding of Argentinean lamb bones to pigs, involving the slaughter of 400,000 animals [Anon, 1969]. And more recently an epidemic in 2001, thought to have arisen from feeding of catering waste (possibly derived from illegally imported meat) to

pigs, with 2030 farms classified as infected and involving the culling of approximately 4 million animals [Anderson, 2002]. The estimated direct costs to agriculture and food production were £ 3.1 billion. The indirect losses to tourism and other industries were estimated at £ 2.7 - 3.2 billion[Francis et al., 2002].

In August 2007 the accidental release of foot-and-mouth disease contaminated material from a research laboratory and vaccine manufacturing plant in Surrey caused the UK's most recent foot-and-mouth disease outbreak. A total of 8 infected premises were identified between the 2^{nd} of August and the 30^{th} of September. After widespread serological testing of stock the EU permitted trade from 31 st December and the World Animal Health Organisation (OIE) declared the UK as once more free of foot-and-mouth disease on the 22^{nd} of February [Anderson, 2008].

1.2.2 Virology

Foot-and-mouth disease is a highly contagious, usually non-fatal disease of cloven-hoofed animals. The disease is caused by a non-enveloped virus of the family *Picornaviridae*. These are single-stranded, positive sense, RNA viruses, about 30nm diameter. It is of the Apthovirus genus as it is acid labile. It exists in seven distinct serotypes: O, A, C, South African Territories (SAT) 1, SAT 2, SAT 3 and Asia 1, most of them with many subtypes [Kitching et al., 1989]. It is also subdivided into topo-types e.g. foot-and-mouth disease virus Type O has eight topo-types. As an RNA virus foot-and-mouth disease virus mutates rapidly. A study of the UK 2001 epidemic estimated a mean 4.3 nucleotide base changes between each separate infected farm [Cottam et al., 2008a]. This property has been exploited in the molecular epidemiology of the disease whereby a substitution history can be used to connect sampled virus in an epidemic tree [Bastos et al., 2003; Bronsvoort et al., 2004; Cottam et al., 2008a,b]. Infection or vaccination with one serotype, or in some cases even a different sub-type of the same serotype, does not confer immunity against another [Mahy, 2005]. The different foot-and-mouth disease serotypes/topo-types have differing transmission characteristics and host susceptibilities; the widely distributed, Pan Asian O topo-types responsible for the UK 2001 epidemic being considered 'aggressive' and difficult to control by vaccination [Mahy, 2005].

1.2.2.1 Virus Properties

The foot-and-mouth disease virus is relatively temperature and pH sensitive. It survives for up to one year at 4^oC but only for 10 days at 37^oC. It is rapidly inactivated by pH of less than 6 or above 9 making disinfection with citric acid or caustic soda convenient and economical bio-security measures [AVIS Consortium, 2002]. Foot-and-mouth disease virus is resistant to iodophores, quaternary ammonium compounds, hypo-chlorite and phenol, especially in the presence of organic matter. Survival times in the environment are largely dependant on moisture content of the medium; the virus will survive in slurry for up to 6 months but only for 14 days in dry faeces. Survival times in soil range from 3 days in the summer to 28 in the winter [Alexandersen et al., 2003].

1.2.3 Hosts

Foot-and-mouth disease infects a wide range of domesticated and wild species. Economically important domestic species include cattle, sheep, goats, pigs and buffalo. Wildlife species include deer, hedgehogs and rats. A complete list is available from the AVIS Consortium website [AVIS Consortium, 2002]. Additionally there are anecdotal reports of a very small number of reported cases of foot-and-mouth disease infecting humans with claims that two people died from the disease in the UK in 1884 [Jeffery, 2001]. Generally the disease is considered to be of minor direct consequence to human health [National Health Service, 2008].

1.2.4 Clinical Illness

The incubation period is 2-14 days from inoculation to first appearance of clinical signs. These signs vary in severity between species and with serotype of virus [Alexandersen et al., 2003] and [Mahy, 2005]:

Signs in Cattle

Foot-and-mouth disease in cattle is normally clinically overt. Signs include pyrexia, anorexia, in-appetence, shivering, reduced milk production for 2-3 days, then smacking of the lips, grinding of the teeth, drooling, lameness, stamping or kicking of the feet

caused by vesicles on buccal and nasal mucous membranes and/or between the claws and coronary band. After 24 hours, these vesicles rupture leaving erosive lesions. Vesicles may also appear on the udder. Recovery from clinical signs is normally within 8-15 days. Complications include tongue erosions, secondary infection of lesions, hoof deformation, mastitis and permanent reduction of milk production, permanent weight-loss and abortion. Young animals may die due to myocarditis with mortality rates as high as 30%.

Signs in Sheep and goats

Clinical signs in sheep are more subtle than in pigs and cattle. Vesicular lesions are less pronounced. Foot lesions may even go unrecognised. Lesions may be evident in the dental pad. Milk yields may be reduced and in a milking flock this may be identified by stockmen. Foot-and-mouth disease may present as sudden death in lambs or as a reduction in fertility rate or fecundity.

Signs in Pigs

Vesicles may be seen on the snout, lips gums and tongue. Vesicles on the feet rupture after 1-2 days leaving eroded lesion and pronounced lameness. Lesions may also be seen on other pressure points such as the knees and hocks. Digits may also slough exposing the laminae underneath. In breeding herds there may be a high mortality rate in piglets with postmortem identification of multi-focal cardiac lesions.

1.2.5 Clinical diagnosis of foot-and-mouth disease

Clinical diagnosis of foot and mouth disease will be based on an exclusion of possible differential diagnoses (see table 1.1 for differential diagnoses [OIE, 2002]) combined with clinical pathology on either blood, milk or lesion samples. In a normally disease-free region, such as the UK, the primary clinician examining an animal presenting with foot-and-mouth disease signs is required to promptly contact the government animal health service in order to expedite diagnosis and application of control measures.

1.2.6 Epidemiology

1.2.6.1 Transmission

Virus can be isolated in all excretions and secretions from an infected animal. These include: Saliva, nasal discharge, vesicular fluid from the mouth and feet, samples of exhaled breath, milk, semen, urine, faeces, vaginal discharge and aborted foetuses. Virus excretion rates vary by species and route. Excretion rates in exhaled breath have been measured at up to 160 virus particles per minute in cattle and sheep and up to 250,000 particles per minute in pigs. Virus excretion is estimated to last from 5 to 11 days with the salivary route of greatest duration [AVIS Consortium, 2002]. Beyond the 11 days, there are reports of virus being isolated from oesophageal (probang) samples for considerable periods after infection of 9 months in sheep, 2.5 years in cattle and more than 5 years in buffalo [Alexandersen et al., 2002a].

1.2.6.2 Animals and Animal Products

Foot-and-mouth disease may be introduced into a previously disease-free region by import of infected or carrier animals [Martinez-Lupez et al., 2008]. It may also be transmitted by animal products; meat from viraemic animals can contain virus particles and be infectious. The increase in pH in meat during the rigor mortis process will normally be sufficient to inactivate the virus. However, this change does not occur in lymph nodes and bone marrow which can remain infectious for several months post-mortem.

Clinically Indistinguishable	**Other Differential Diagnosis**
Vesicular Stomatitis	Mucosal Disease
Swine Vesicular Disease	Infectious Bovine Rhinotracheitis
Vesicular Exanthema of Swine	Bluetongue
	Bovine Mammillitis
	Bovine Papular Stomatitis
	Bovine Viral Diarrhoea
	Rinderpest

Table 1.1: Diseases with a similar clinical appearance to foot-and-mouth disease

1.2.6.3 Transmission by Personnel, Equipment and Vehicles

Foot-and-mouth disease virus remains infectious in faecal fomites and fluid droplets that may become attached to personnel, vehicles or equipment. This is particularly relevant in the case of veterinary surgeons, farm workers and delivery services that may travel from farm to farm unwittingly transmitting infection. These routes pose particular problems during disease epidemic control as in the absence of detailed records they are difficult to trace. Thrusfield et al. [2005b] documents putative sources of infection for farms in the Dumfries and Galloway region during the UK 2001 foot-and-mouth disease epidemic, with personnel movements estimated to be the dominant cause of transmission events. These estimates were based on collated data of potential epidemiological links and timings of infection on associated premises.

1.2.6.4 Airborne Spread

Foot-and-mouth disease virus exhaled by animals may be spread as an aerosolised viral plume from one group of animals to another. Depending on the virus serotype and the immediate meteorological conditions viral plumes may be tens of kilometres long. In 1981 an outbreak of foot-and-mouth disease in cattle on the Isle of Wight was attributed to airborne transmission from a pig unit in Henansal, Brittany [Gloster et al., 1982]. Optimal temperature conditions with high relative humidity and a 7 m/sec southerly wind were thought to result in a long viral plume and hence successful transmission. This and other airborne transmission events have been modelled by Sorensen et al. [2000]. Airborne transmission was also considered to be the cause of the rapid initial escalation of the 1967 epidemic and also the route from the index case pig farm in 2001 to a neighbouring sheep unit from which movement of animals seeded large areas of the UK [AVIS Consortium, 2002; DEFRA, 2008].

1.2.6.5 Transmission by Wild Birds and mammals

The Northumberland Report on the 1967/68 epidemic reported evidence that birds may carry foot-and-mouth disease virus for up to 91 hours [Anon, 1969] although there are no epidemiological studies available to assess the importance of this transmission mode.

Species	TCID50 respiratory	TCID50 oral	Excretion (estimated virus units day^{-1})
Cattle	10 - 1000	3×10^6	10^5
Sheep	15 - 100	Not known	10^5
Pigs	400	105	10^8

Table 1.2: Foot-and-mouth disease virus susceptibility (Tissue Culture Infectious Dose causing infection in 50% of individuals – TCID50) and airborne excretion

1.2.6.6 Susceptibility

Foot-and-mouth disease can be acquired by inhalation, ingestion or through damaged epithelium. Approximate infectious doses vary by route and by species; they are shown in table 1.2 alongside estimates of airborne excretion. These figures are approximate as they are experimental and vary by serotype [AVIS Consortium, 2002].

1.2.7 Carriers

A foot-and-mouth disease carrier is defined as an animal from which foot-and-mouth disease virus can be recovered from the oropharynx at 28 days or later, post-infection[Sutmoller et al., 1968]. Virus has been found after this period in the saliva of cattle with the site of virus persistence identified as the dorsal surface of the soft palate [Murphy et al., 1994]. In sheep the virus persists in the tonsillar area [Burrows, 1968]. Although virus persistence in these areas is a normal feature of infection the probability of an animal being a carrier decreases with time and was also related to the severity of the original challenge infection [Hedger, 1970; Moonen et al., 2004]. The significance of carrier animals in the transmission of foot-and-mouth disease and implications in its control is contentious. Definitive transmission from carriers to cattle has been demonstrated in buffalo [Dawe et al., 1994], has been anecdotally reported in cattle and is considered unlikely in sheep [Alexandersen et al., 2002a].

1.2.8 Vaccination

The first attempts to develop a foot-and-mouth disease vaccine were in the 1920's with work concentrating on attenuating the virus to produce a live, non-pathogenic vaccine. Much later, in the 1960's, this work was continued but virulence of the resulting vaccine rendered the vaccines practically unusable. Live vaccines may be produced

using current vaccine technology however they would have the disadvantage, compared with inactivated vaccines, that vaccinated animals would be hard to differentiate from infected animals. The first effective vaccines were developed by Waldermann et al [Desmettre, 1995] and were inactivated vaccines manufactured by extracting fluid from epithelium and tongue vesicles from deliberately infected cattle and inactivating it with formaldehyde and using an aluminium hydroxide adjuvant [Doel, 2003]. The inconvenience of using an infected animal host was removed by Frenkel in 1947 who grew foot-and-mouth disease virus in a culture medium derived from tongue epithelium of healthy, slaughtered cattle. This method was used to manufacture foot-and-mouth disease vaccine until 1962 when immortalised cell line culture methods were developed by various researchers using hamster kidney cells as a growth medium [Capstick et al., 1965; Doel, 2003]. These immortalised cell line methods reduce logistic and sterility-testing problems compared to culture in epithelia media. Vaccine quality was further improved by the introduction of more effective inactivation technologies including use of aziridine acetylethyleneimine [Brown and Crick, 1959] and now binary ethyleneimine (BEI) [Bahnemann, 1975]. BEI inactivation is now required by regulatory authorities to achieve active virus particle contamination rates less than one per 10,000 litres of product [OIE, 2009a]. The inactivated products are currently concentrated and purified by ultrafiltration. As the vaccine is inactivated, an adjuvant is required to stimulate the immune response. Saponin is now added to the aluminium hydroxide adjuvant for use in ruminants. Oil based adjuvants are typically used in pig vaccines [Intervet, 2005]. The OIE Manual of Diagnostic Tests and Vaccines for Terrestrial Animals contains detailed specifications regarding the manufacture, efficacy, safety and testing of foot-and-mouth disease vaccines [OIE, 2009a]

1.2.9 Geographical Distribution

Foot-and-mouth disease is endemic in parts of Asia, Africa, the Middle East and South America with sporadic outbreaks in many disease-free areas. North America has been disease-free since 1929 when the last outbreak occurred in California [Segarra and Rawson, 2001], Australia has been free since 1872 [Cannon and Garner, 1999] and New Zealand has no history of any foot-and-mouth disease outbreaks [Thomson, 2009]. See figure 1.1 for world distribution of foot-and-mouth disease outbreaks and endemic areas in August 2009 [OIE-WAHID, 2008].

Geographical distribution is serotype related with O and A serotypes of foot-and-mouth disease widely distributed and the SAT serotypes found mainly in Africa. Asia 1 serotype is found only in Asia and the C serotype is extinct.

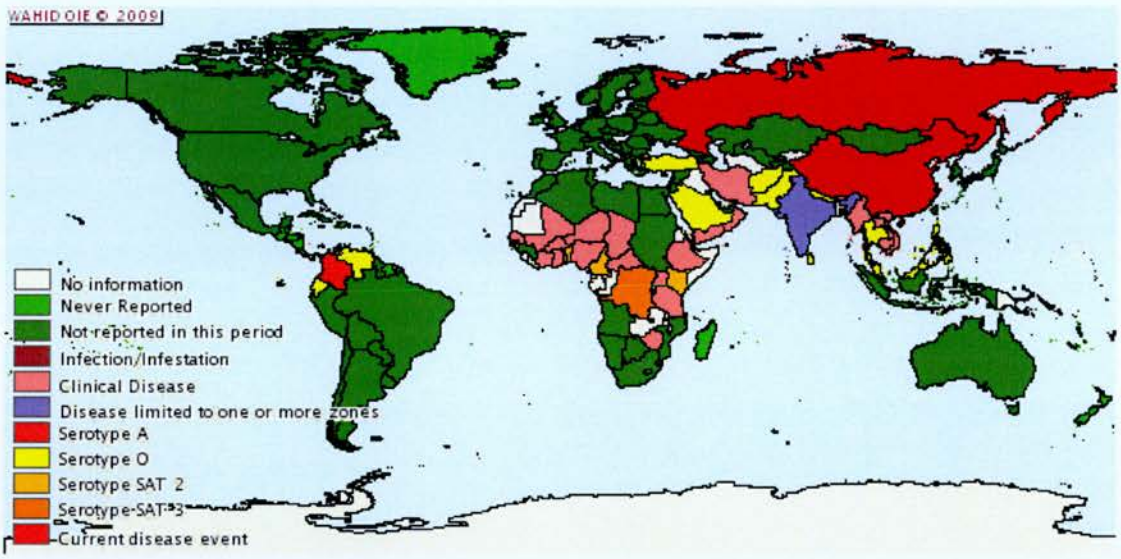


Figure 1.1: Distribution of foot-and-mouth disease outbreaks and endemic areas from OIE WAHID reporting system to August 2009.

1.2.10 Diagnosis

The initial diagnosis of foot-and-mouth disease will normally be based on clinical signs. This is the main mechanism used for initial detection of an outbreak. It relies on examination by stockholders, clinical veterinarians, animal haulers, abattoir staff and official veterinary surgeons at abattoirs and ports of entry. Clinical diagnosis is inexact and its performance varies with species. Sheep and goats, particularly, may harbour foot-and-mouth disease without showing overt clinical signs [Mahy, 2005]. To confirm clinical diagnosis or as part of control and surveillance operation laboratory based tests will be used as they will generally increase the sensitivity and specificity of the overall diagnostic process. Laboratory tests for foot-and-mouth disease are carried out at approved laboratories with biological security precautions in place owing to the potential for release of infectious material from submitted samples. The UK laboratory for such diagnosis is the Institute of Animal Health at Pirbright.

Tests can be based on detection of either virus antigen or antibodies raised to the virus by the host animal and are summarised here:

Antigen Based Tests

Virus isolation in tissue culture Samples are cultured in a cellular growth medium and cytopathic effects observed directly.

ELISA antigen test Samples are incubated with anti-foot-and-mouth disease virus antibodies which may be directly or indirectly labelled with an enzyme which is then used to promote a reaction with a conjugate. The change in conjugate is normally detected colourimetrically.

PCR - polymerase chain reaction Viral RNA is amplified using primers and polymerase enzymes. Results of the amplification process are normally detected using electrophoresis.

Antibody based structural protein serological tests

Virus Neutralisation Test Serum is incubated with a known sample of tissue pathogenic foot-and-mouth disease virus. The result is then tested using tissue culture to determine if neutralising antibodies in the serum have prevented cytopathic activity in the virus.

Antibody ELISA Similar to the ELISA antigen tests, a known foot-and-mouth disease virus antigen is used to bind potential foot-and-mouth disease antibodies in the sample and detected using an antibody linked enzyme / conjugate system.

Non-Structural Protein Serological Tests — for NSPs

These tests identify antibodies to non-structural proteins (discussed later) that are present in animals in which foot-and-mouth disease virus has replicated but are not present in non-exposed, foot-and-mouth disease vaccinated animals.

ELISA A number of ELISA based assays have been developed using some or all of the non-structural 3ABC proteins. These test are reviewed by Brocchi et al. [2006]

Enzyme linked immuno-transfer blotting Also called Western blotting, is an electrophoresis based technique where proteins from the sample are separated on a gel under an electric field. They are then transferred to a membrane and identified

using specific linked antigen/antibodies. The EITB technique is claimed to have high specificity and be suitable as a confirmatory test for foot-and-mouth disease virus non-structural proteins [Bergmann et al., 1996].

1.2.11 Diagnostic tests in post epidemic surveillance

Post-epidemic serological surveillance for foot-and-mouth disease is currently focused on the use of enzyme linked immunosorbent assays (ELISAs [Lequin, 2005]). These tests can identify antigen or antibody to foot-and-mouth disease. The overall methodology relies on coated wells in which antigen or antibody is trapped. An enzyme labelled antibody then promotes the conversion of a substrate to an optically detectable product (either by optical absorbance or fluorescence). The methodology may be modified, for example, by having foot-and-mouth disease virus antibodies in the sample compete with labelled antibodies where a sample *low* in antibody will result in *higher* proportions of bound, labelled antibodies and hence a higher optical absorbance reading.

The standard ELISA test used in the UK 2001 foot-and-mouth disease epidemic for the detection of foot-and-mouth disease antibodies in unvaccinated cattle is the cELISA described in detail in Paiba et al. [2004]. In this test, foot-and-mouth disease antigen is fixed to antibody coated wells. Then the test serum is added together with a competing guinea pig anti-foot-and-mouth disease antibody. After incubation the wells are washed and depending on the relative concentrations of antibodies in the test and competitive sera there will be a similar proportion of antibodies bound to the antigen in the wells. The proportion of test antibody bound is assayed by using an horseradish peroxidase labelled anti-guinea pig immunoglobulin. This binds to the competitive antibodies and is measured by final incubation with a substrate. The resulting optical density is related to the proportion of bound competitive antibody i.e. the greater the optical density the lower the antibody concentration in the test serum. The optical density result is corrected by calculation of a percentage inhibition (PI) using the following formula :

$$PI = 100 - \frac{OD_{sample} - OD_{conjugate}}{OD_{control}} \times 100$$

where OD is the optical density of the test well and $OD_{control}$ is the optical density of the control wells with *no* test serum and $OD_{conjugate}$ is the optical density of the control

wells with no test serum or competitive serum (i.e. background conjugate effect). (note: formulae for PI will vary from test to test in detail e.g. some methods do not subtract $OD_{conjugate}$ from the OD_{sample})

The cELISA test is serotype specific and also has the disadvantage that it cannot be used in vaccinated animals as foot-and-mouth disease vaccination elicits an antibody response against the structural proteins detected by the cELISA. More recent test developments have produced ELISA tests that detect non-structural proteins (NSPs) of foot-and-mouth disease virus. Non-structural proteins are proteins that are coded for by the foot-and-mouth disease virus but are not part of the virus capsid. They are classified as protein products from the genes 3A, 3B, 3C and 3D (see figure 1.2). Hence it is possible to develop an inactivated vaccine that will elicit immunity to the surface capsid proteins of the virus but will not elicit antibodies to NSPs since there will have been no viral replication in the host animal. However if an animal becomes infected with foot-and-mouth disease NSPs will be produced during virus multiplication and are released during cell lysis. These NSPs elicit an immune response with the formation of antibodies to NSP. Thus previously foot-and-mouth disease infected animals can be detected by the presence of NSP antibodies irrespective of their vaccination status.

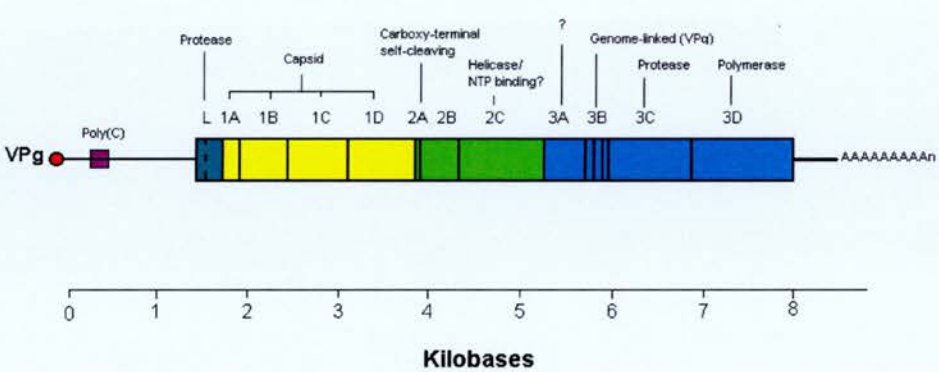


Figure 1.2: Foot-and-mouth disease genome (from Information System for the OIE/-FAO FMD Reference Laboratories Network)

The major NSP specific ELISA tests are reviewed and evaluated in Brocchi et al. [2006]. Here I will describe the outline protocol for the CEDITEST NSP test as this is the main test used in the Turkish sero-surveillance presented in chapters 3, 4 and 5. The assay is described in detail in Sorensen et al. [2005]. In summary:

Plain ELISA plates are coated with a monoclonal IgG antibody against NSPs 3ABC. These plates are then incubated with 3ABC antigen (manufactured in baculovirus) to give a plate coated with IgG antibody with a bound layer of 3ABC antigen. Serum or control samples are then added and the plates are incubated. Test serum containing 3ABC antibodies will bind to the antigen layer forming an antibody-antigen-antibody sandwich. Then horseradish peroxidase labelled monoclonal IgG antibody is incubated with the plates. This antibody will bind with otherwise unbound antigen. After final washing the amount of bound, labelled antibody is assessed by addition of a substrate and colorimetery at 450nm. The percentage inhibition result is calculated from a corrected optical density using OD_{NIR} — a non-inhibited reference with buffer substituted for test samples. A high percentage inhibition signifies low optical density i.e. greater binding of serum IgG and hence a positive result. The manufacturer recommended cut-off is a PI of 50%.

$$PI = 100 - \frac{OD_{sample}}{OD_{control}} \times 100$$

1.3 Economics of foot-and-mouth disease

Foot-and-mouth disease is considered to be of high economic importance in the agricultural systems of the developed world [Prempeh et al., 2001]. This arises from its effects on animal production at an individual level and, depending on consequences to trade and control, from its contagious nature, high infectivity and significant morbidity.

1.3.0.1 Economics at a Local Level

Foot-and-mouth disease has an impact at a local farm or producer level influencing the productivity of agricultural enterprises [Doel, 2003]. It has an impact on many aspects of livestock production systems as listed below causing ...

- an average 25% reduction in annual milk yield of dairy cows.
- meat production to be reduced with some reports estimating a 10-20% increase in time to maturity and a 20% reduction in annual pig meat production in infected animals.

- animal draught power to be reduced by 60-70% in the first month after infection.
- fertility rate to be reduced by up to 10% due to increased abortion rate.
- very high mortality losses in young animals of up to 30%.
- culling loss of chronically infected animals
- gross disruption of farming practices

A further local economic loss is that of missed selection opportunity. It is unlikely that the performance of a chronically infected animal will exactly correlate to its performance if it were healthy. Hence selective breeding pressures applied to an animal population with transiently endemic foot-and-mouth disease will results in a herd that may not perform optimally if foot-and-mouth disease were not present. This difference is termed the "lost selection benefit" and represents an opportunity cost of disease.

1.3.0.2 Effects at a National Level

The national and international economic consequences of foot-and-mouth disease arise from control measures generally enacted by statutory disease control regulations, emergency control legislation and self-imposed control and bio-security measures. The highly infectious nature of the disease makes movement restriction of livestock and temporary closure of livestock markets prime components of control strategy [Doel, 2003, page 133]. However, these restrictions effectively cease livestock trade for the duration of the control period. These regulations and compensation systems for farmers' losses during an outbreak will shift the costs of foot-and-mouth disease losses from the farms experiencing the loss to the general agricultural economy, other sectors and the state. This can make accurate estimation of the economic consequences of an foot-and-mouth disease outbreak difficult.

1.3.1 Economic Consequences of 2001 foot-and-mouth disease Epidemic

The state costs to the UK exchequer of controlling the 2001 foot-and-mouth disease epidemic are estimated within the National Audit Office (NAO) 2002 report [National Audit Office, 2002] as £3.1 billion in total. This is represented by £1.4 billion in compensation to farmers for culled stock and £ 1.7 billion as other disease control measures.

The report estimates the total loss to the rural economy as more than £ 5 billion arising mainly from tourism losses. £355 million of this is uncompensated losses by agricultural producers (about 20% of 2001 farming income). The NAO report that the food industry cost, downstream of agriculture, was around £ 115 million. The NAO report estimates that the overall impact on the UK economy was modest due to substitute expenditure by consumers; the epidemic reducing UK GDP by 0.2% (of total GDP which is approximately £1.8 trillion i.e. a £3.6 billion cost) [DEFRA, 2008; Francis et al., 2002].

DEFRA report the medium-term effects of the 2001 foot-and-mouth disease epidemic on UK agriculture [Francis et al., 2002]. In summary: It was found that a small proportion of farmers from infected premises (6%) intended to leave agriculture. Farmers expected a further reduction in number of people employed - this continued the pre-epidemic trend. A 2% increase in the use of contractors and share farming was expected. Restocking was planned in about 70% of culled premises but to lower, on average, levels than before the epidemic. The number of cattle in the dairy sector was expected to fall but with production levels maintained due to increased efficiency. There was also evidence of an increasing diversification of farming enterprise into organic and environmental management schemes.

1.3.1.1 Foot-and-mouth disease Control

The definition of *under control* was one of the contentious issues surrounding the 2001 UK foot-and-mouth disease epidemic [BBC News Online, 2001]. For the purposes of this section, disease control is considered to be any intervention that is enacted with the intention to reduce epidemic size, duration or severity in its broadest sense. Foot-and-mouth disease control can adopt two, non-exclusive approaches; either to reduce the probability that units become infected or to reduce the probability that infected units transmit disease. These actions can be applied within the context of the unit being an individual, an animal group/farm or a region/nation.

1.3.1.2 Import Control and Stock Movement

Following the first recorded appearance of foot-and-mouth disease in the UK in 1839 the disease was assumed to be un-preventable and there is no record of specific control

measures [Anderson, 2002; DEFRA, 2008; Woods, 2004]. In fact, prior to 1841 animal importation was prohibited but this restriction was removed in 1842 subject to payment of the requisite, per head, duty. This duty was later removed in 1846 by Sir Robert Peel, once more permitting free trade. However, a UK Rinderpest epidemic 1864-66 motivated widespread movement and import control legislation. In addition to this, a weekly report was published in the London Gazette detailing contagious disease outbreaks. These controls were viewed as important in the elimination of Rinderpest and were legislatively fine-tuned to help control foot-and-mouth disease. Import restrictions persist to date with a balance being sought between the desires to maintain international free trade and the wish to protect the health of the national herd. The need to recognise this balance is encapsulated in the World Trade Organisation Agreement on the Application of Sanitary and Phyto-sanitary Measures [World Trade Organisation, 2008]. This outlines a basis by which nations can protect their animal (and human) populations against disease whilst aiming to prevent the use of disease control arguments as a mechanism to disrupt free trade [World Trade Organisation, 2008]. The EU maintains a list of countries outside the community which are permitted to export animal products into the EU. This list will be revised from time to time with changes in the international disease status and changes in animal disease science [European Commission, 2006]. Such imports have to pass through approved Border Inspection Posts. In the UK these are maintained by the State Veterinary Service [Anderson, 2002].

1.4 Surveillance for disease-freedom

Surveillance in the context of animal disease is a process where individuals and/or the population are observed in order to assess disease status with an intention to take action if the status changes [Salmon, 2003]. In the specific case of surveillance to demonstrate disease-free status the methodology and consequences of such surveillance will be informed by national and international codes and regulations.

1.4.1 Definitions of Disease-Freedom

The OIE in the Terrestrial Animal Health Code - 2005 [OIE, 2005b] define a disease-free zone as:

> ...a zone in which the absence of the disease under consideration has been demonstrated by the requirements specified in this Terrestrial Code for free status being met. Within the zone and at its borders, appropriate official veterinary control is effectively applied for animals and animal products, and their transportation.[OIE, 2005a]

The OIE Terrestrial Code lays down requirements for disease-free status for countries and zones within countries for control with and without vaccination. The requirements for a country to obtain disease-free status require the following requirements to be met depending on the class of disease-free status required as follows:

Disease-freedom in the absence of vaccination

Foot-and-mouth disease may be controlled without vaccination by application of strict import and biosecurity measures. To be classified as foot-and-mouth disease-free without vaccination a country will need to provide evidence to the OIE of regular and prompt disease reporting and a declaration of:

1. No outbreak in the previous 12 months.
2. No evidence of infection in the previous 12 months.
3. No vaccination in the previous 12 months.
4. Supply evidence of surveillance for foot-and-mouth disease and virus in accordance with OIE codes.
5. No importation of foot-and-mouth disease vaccinated animals since the country's own cessation of vaccination.

Disease-freedom with vaccination

Where importation controls are difficult it is possible for a country to maintain a lower standard of freedom with vaccination included as a control measure. Achieving this status requires:

1. Evidence of regular and prompt disease reporting.
2. Declaration of no outbreak in the previous 2 years and no evidence of foot-and-mouth disease circulation in the previous 12 months.
3. Documented evidence that:
   (a) Surveillance for FMD and FMDV circulation in accordance with OIE codes is in operation, and that regulatory measures for the prevention and control of FMD have been implemented;
   (b) Routine vaccination is carried out for the purpose of the prevention of foot-and-mouth disease;
   (c) With a vaccine that complies with the standards described in the OIE manual.

If a foot-and-mouth disease-free country, where vaccination is practised, wishes to change its status to a foot-and-mouth disease-free country where vaccination is not practised, the country should wait for 12 months after vaccination has ceased and provide evidence showing that foot-and-mouth disease virus circulation has not occurred during that period. The OIE provide further definitions referring to foot-and-mouth disease-free zones with and without vaccination control. These parallel the foot-and-mouth disease-free country regulations with additional requirements for physical, geographical or buffer zone areas to separate disease-free areas from other areas. After a foot-and-mouth disease epidemic, specific OIE regulations refer to the process by which a formally disease-free country may regain its disease-free status of either free with or without vaccination [OIE, 2009b].

1.4.2 History of freedom from disease surveillance design

Freedom from disease surveys are part of the mechanism by which a zone is demonstrated to be free of a nominated disease. Their most common application is after an outbreak or epidemic. A number of samples are taken from animals in the zone, they are biologically tested for evidence of disease and after suitable statistical analysis the survey result will be reported. Classically, a successful result is of the form:

> The survey has found no evidence of disease. If disease were present in at least $x\%$ of epidemiological units and within those infected epidemiological units if it were present in at least $y\%$ of animals the survey would detect disease with a $z\%$ confidence.

Where x, y and z define the statistical performance and epidemiological design criteria of the survey.

Hence these surveys take the form of a frequentist statistical test with a null hypothesis of, say, disease being present and a size or Type I error of, for example, $(1 - z)$ i.e. $\alpha = (1 - z)$ [Cameron and Baldock, 1998b]. The development of the design and analysis of these surveys is reviewed in Cameron [1997] and in Baldock [1998]. The following represents a summary together with outlines of more recent issues and developments.

The design of such surveys regarding sample sizes has undergone a process of improvement using increasingly sophisticated models for the disease, survey and diagnostic test process. The original sample size calculations assumed sampling from an infinite population using a perfect diagnostic test with a disease that was reliably present at a design prevalence [Cameron, 1997]. Sophistication in survey design now includes finite population models (using hypergeometric models of within herd sampling) and imperfect knowledge of imperfect diagnostic tests [Hanson et al., 2003; Audige et al., 2001]. Generally this will increase the required sample size to declare disease-freedom with a given confidence level. The later approaches are also not solvable analytically so require numerical solutions to determine approximate sample sizes. Public domain software exists to provide these solutions such as FreeCalc 2 [Cameron and Baldock, 1998b].

Further work by Donald et al. [1994] challenges the regulatory standards and surveillance designs that make the assumption that if disease is present within a herd that it will be present at a minimum within-herd prevalence. In reality a herd may contain as few as one diseased animal if the herd has only recently become infected, if there is limited susceptibility as in the case of vaccination or if there were segregation between herd members. Even if the mean prevalence across all infected herds meets some design level, heterogeneity across herds will reduce the overall efficacy of the surveillance process so it may fall below the design level. Donald et al. [1994] discusses and models this

although the results are not extended to demonstrate the consequences in a real-world, two stage disease survey.

A particular problem of survey design and analysis occurs in small animal groups. Many regulatory codes require a survey to be able to detect disease in an epidemiological group (i.e. herd or flock) at a certain confidence if it is present at a certain prevalence as described above. With moderately insensitive diagnostic tests this can be achieved in most groups by sampling sufficient animals. However in small groups even sampling *all* the animals may not provide the required group level sensitivity/confidence. This issue has been discussed by Martin et al. [1992] and in detail by Greiner and Dekker [2005] who offer models to estimate survey performance in small herds and possible solutions including consideration of *all* small herds as infected or the alternative suggestion that small herds pose a less important disease threat and may be considered uninfected.

An alternative to regulation-informed design is presented by Hohle and Jorgensen [2003] regarding prevalence estimates of disease in a herd. Their study uses a loss function for sensitivity and specificity to develop an analytic function for utility and hence explore sensitivity to uncertainty in test parameters. The analysis does not extend to multi-herd scenarios and does not consider explicitly the disease-freedom issue.

1.4.2.1 Bayesian approaches to survey analysis and design

Most previous work addresses the issue of survey design and analysis as a frequentist hypothesis test. A Bayesian approach views probability as a measure of subjective belief allowing epidemiologists and stake-holders to refer to "probabilities of disease-freedom". Johnson et al. [2004] discuss a Bayesian approach to sample size calculation within a single stage (herd/flock) survey framework and illustrate it using foot-and-mouth disease screening within a herd. Suess et al. [2002] derive a Bayesian model for disease-freedom estimation illustrated with examples using Newcastle Disease and Porcine Reproductive and Respiratory Syndrome. In a further extension of this approach Choi et al. [2006] describe a Bayesian approach to use the full continuous information from diagnostic tests rather than dichotomising the results with a cut-off (positive/negative) value for each animal. The study illustrates the estimation of probability of infection for each individual in a sampled group. Bayesian approaches have the advantage that they allow the incorporation of prior information into the analysis, often provide a naturalistic framework for specifying the problem and produce probabilistic statements about

the estimated quantities. In the context of testing for disease-freedom these probabilistic outputs can be used to provide direct input into economic and utility based models. Critics of Bayesian approaches highlight the need for, potentially subjective, prior information arguing that this makes the whole process subjective. Another issue is that Bayesian approaches will often require large amount of numerical computation increasing costs and delaying results.

1.4.2.2 Serial and parallel testing

Many disease-freedom surveys will use an initial, screening, test on each sampled animal. These tests are selected considering diagnostic performance but also speed and economy. Often the diagnostic test will have an imperfect specificity - it will give false positive results. To reduce the impact of these results, further tests can be done on animals testing positive using a more specific, confirmatory, test. Alternatively, in situations where no single test is sufficiently sensitive, another test can be done in parallel to increase the probability of detecting diseased animals. Gardner et al. [2000] discuss these approaches with reference to disease-freedom estimation and considers the impact of conditional dependence between tests, where co-varying tests may give less additional information than conditionally independent tests.

1.4.2.3 Retrospective evaluation of surveillance designs

The performance of a given survey design may be retrospectively assessed in simple cases by analytic methods. In more complex settings, with uncertainty of diagnostic test performance and epidemiological parameters, simulation models will be required. In Tsutsui et al. [2003] a stochastic model was used for evaluation of the foot-and-mouth disease surveillance in Miyazaki, Japan after the outbreak in 2000. The study models within herd prevalence using Reed-Frost model [Carpenter, 1984]. It incorporates uncertainty in diagnostic test performance and estimated herd level sensitivity at 71 to 76.5%. Similarly, Carpenter and Gardner [1996] use simulation modelling to explore interpretation of a serologic survey for porcine parvovirus. The effects of diagnostic test performance and number of animals tested are investigated demonstrating that extrapolation from animal-level test performance to herd level survey performance is

complex and non-linear. The study did not explicitly consider disease-freedom as such but classified herds into sero-prevalence bands.

1.4.2.4 Use of multiple information sources — scenario trees

In practical applications, disease-freedom surveys using serological methods will not be used in isolation; the results of serological surveillance will be interpreted in the light of previous surveys and clinical information. In a review of disease-freedom approaches this is discussed by Cannon [2002]. His work also highlights that most studies consider the technical issues of surveys to demonstrate disease-freedom rather than standing back and considering what these surveys could most usefully estimate or detect. Cannon [2002] introduces a points system for scoring tests performed in the face of estimated disease prevalence incorporating test sensitivity, number of tests performed and expected disease prevalence and discusses combination of prevalence from multiple surveys. The approach appears to be based on point estimates and probability theory and does not consider the correlation of survey estimates. The use of multiple surveys is further expanded in Martin et al. [2007b,a] where the "scenario tree" methodology is described. This uses Bayes theorem (see later) to update a probabilistic inference from one surveillance activity with the results of another [Salmon, 2003, page 172-174].

1.4.2.5 The influence of vaccination on surveillance

Paton et al. [2004] reports EU Council Directive 203/85/EC which gives provision for vaccination control of foot-and-mouth disease and defines post-vaccination sero-surveillance requirements. They discuss issues of test sensitivity and specificity in a vaccinated population. Sutmoller and Gaggero [1965] (cited by Paton et al. [2004]) observed a 50% within herd prevalence four months after a vaccine trial failure in Brazil. Paton et al. [2004] also reports a recent trial where 45% of 20 vaccinates became infected 21 days after exposure in a research setting. The issues of herd false positives and small herds are further discussed with reference to the work of Greiner and Dekker [2005] with the suggestion that positive herds are re-sampled and re-tested to demonstrate absence of *circulation* rather than *infection*. Later, Arnold et al. [2007] developed a model to estimate the residual carrier population after a foot-and-mouth disease outbreak to assess the efficacy of non-structural protein based testing. Their

study estimated that only 0.2% of herds in the UK would contain carrier animals after an outbreak and that, within these herds, the number of carriers would be low. On this basis they suggest testing of all vaccinated herds and culling of just the test positive animals. This approach assumes that the animal-level tests are adequate in terms of both sensitivity and specificity to avoid missing carrier animals and excessive culling of healthy stock.

1.4.2.6 Risk based surveillance

Recently attention has been paid to development of risk based strategies for the estimation of population parameters and demonstration of disease freedom. Williams et al. [2009b] discusses targeted techniques for estimation of population parameters and Williams et al. [2009a] introduces Poisson sampling to demonstrate freedom from disease within a flock of animals and allow estimation of population parameters based on a measurable covariate that is associated with risk of disease. Cannon [2009] discusses the theory behind quarantine inspection and disease surveillance with limited resources proposing analytical expressions to be optimised under different objective criteria for optimisation. In chapter 2 I use a simulation based approach to this problem to estimate the performance of a risk based targeting approach estimating risk from an epidemiological model.

1.5 Bayesian inference

In chapter 5 I use a Bayesian model to estimate prevalences and diagnostic test performance measures based on sero-surveillance data from the Thrace region of Turkey. The Bayesian approach to statistical inference is described in detail elsewhere (Lee [1997] and others). Here I will provide a brief outline.

Bayesian inference derives from Bayes theorem and considers probability to be a measure of subjective belief. The theorem states that the probability of an event, given some observed data, can be calculated from the probability of the event before the data is considered; the likelihood of the event given the data and the probability that the data will be observed:

$$Prob(E|X) = \frac{Prob(X|E) \times Prob(E)}{Prob(X)}$$

Where E is the event and X the data.

For a multi-variable model with continuous parameters this extends to:

$$f(\theta|X) = \frac{f(X|\theta) \times f(\theta)}{f(X)}$$

Where θ is a vector of parameters and $f()$ represents the probability distribution function.

Although it is frequently straightforward to write down an expression for $f(\theta|X)$ only simple problems have closed solutions to allow description of the posterior distribution and summary statistics of θ. Normally, numerical methods are required to sample from the posterior distribution for θ. The JAGS [Plummer, 2003] software used in the Thrace analysis in chapter 5 uses Gibbs Sampling [Lee, 1997, page 259] where draws of values are made randomly from distributions based on only the immediate previous parameter values. This is referred to as Markov chain Monte Carlo (MCMC) methodology as the system has Markov Chain properties (a series where the next value is conditional only on the immediate previous value) and is random hence 'Monte Carlo' after the Monte Carlo Casino in Monaco.

Gibbs sampling is a numerical method for drawing samples from a multivariate distribution. It requires a distribution for each parameter conditional on the other parameters in the posterior. From a starting set of parameter values, samples are drawn sequentially for each parameter conditionally on the remaining current parameter set. Each new draw is then substituted into the current parameter set. In the long term as the number of draws tends to infinity the distribution of drawn values will tend to the posterior distribution of the parameter. For example in a three parameter model of θ at iteration i draws of θ are made sequentially from the conditional distributions:

$$\begin{aligned}
\theta^1_{i+1} &\sim Prob(\theta^1|\theta^2_i, \theta^3_i) \\
\theta^2_{i+1} &\sim Prob(\theta^2|\theta^1_{i+1}, \theta^3_i) \\
\theta^3_{i+1} &\sim Prob(\theta^3|\theta^1_{i+1}, \theta^2_{i+1}) \\
\theta^1_{i+2} &\sim Prob(\theta^1|\theta^2_{i+1}, \theta^3_{i+1})
\end{aligned}$$

The posterior distribution values simulated by this method are only asymptotically distributed as if from the true posterior distribution, i.e. only after an infinite number of samples has been drawn. There will be a finite period during which the samples are highly unrepresentative of the posterior distribution. Identification of the point at which samples can be reasonably considered to approximate the posterior distribution poses a continuing problem for numerical Bayesian inference. A number of approaches exist as discussed in chapter 5 which permit identification of simulation series that have not adequately converged on the posterior distribution. However these methods can only identify *failure to converge*; there are no formal mechanisms by which convergence can be proven. In practical application, examination of multiple long simulation runs, predictions from the model, visual examination of convergence behaviour and absences of non-convergence diagnostics are used to validate MCMC estimation. For a brief, light hearted but enlightening opinion about MCMC diagnostics and their 'bogosity' see Geyer [2005].

The major advantages of Bayesian inference for disease surveillance problems are that the methodology permits complex, naturalistic formulation of models that describe the observed data conditionally on epidemiological and diagnostic test parameters and produces results that capture the uncertainty of the parameter estimates considering the data, model and prior assumptions.

Chapter 2

Risk based surveillance design

2.1 Summary

Post epidemic surveillance will be more efficient if sample selection uses spatial disease models to select farms for sampling and to order their sampling i.e. a higher system sensitivity of infection detection can be achieved for a given sample size.

Current post epidemic sero-surveillance will normally be carried out within the appropriate surveillance zone (10 Km radius for foot-and-mouth disease) and will involve sampling of either all farms (if vaccinated) or a probability sample of farms (if not vaccinated). It is unlikely that an arbitrary sampling strategy will be optimally efficient. A better strategy is to estimate the benefits gained by sampling each farm in the restricted area. This allows the estimation of incremental benefit of sampling each farm and hence allows farms to be sampled in order of decreasing benefit, until the sampling budget is exhausted or surveillance performance targets are met. Furthermore, if farms are sampled in order of decreasing estimated risk of latent infection the time to declaration of disease freedom may be reduced.

In this chapter I adapt a previously published model of the UK 2001 farm foot-and-mouth disease infection risk [Diggle, 2006] to estimate the probability that the farms in Devon had been infected during the outbreak in that region. I then add a model estimating the probability that these farms would have been identified as infected during the epidemic. This allows estimation of the post-epidemic probability of each farm harbouring an undiscovered infection. Adjustment of this result by the individual farm

level diagnostic sensitivity of surveillance estimates the probability that a given farm, if sampled and tested, will result in the detection of a previously undiscovered (latent) foot-and-mouth disease infection. Aggregation of these results for a given sample set of farms allows estimation of the system sensitivity of that set of farms as a tool to detect post-epidemic infection in the region.

The two methods of selecting this sample set of farms are compared across a full range of sampling budgets (from one farm to the entire farm population). This allows comparison of the system sensitivity (ability to detect infection in the area) of random farm selection versus risk model based selection.

Additionally, assuming that samples take a finite time to process from collection to result, the effect of ordering samples by risk is also explored. This estimates the potential of risk-ordered sampling to detect any infected farms earlier than randomly ordered sampling and hence expedite subsequent disease control activities and restoration of disease-free status.

The study showed that it was possible to achieve a system sensitivity (probability of detecting infection in the area if present) of 95% by sampling, on average, 1083 farms with random sampling and 225 farms with risk based sampling. Risk ordering the sampling process resulted in detection of positive farms, if present, 15.6 days sooner than with randomly ordered sampling, assuming 50 farms were tested per day. This would potentially allow declaration of disease freedom and resumption of international trade 15.6 days earlier.

2.2 Introduction

After the apparent end of an animal disease epidemic a country will normally benefit from demonstrating that infection is no longer present in its livestock. This is both to satisfy international trade requirements [OIE, 2009b] and also to identify previously undiscovered infection to prevent recrudescence. Demonstration that disease has been controlled also has domestic societal and political advantages. The process is described as *demonstration of disease freedom* [Cameron, 1997; OIE, 2009b]. With many diseases this process includes a prescribed sampling of livestock in the vicinity of the previously infected premises. Appropriate diagnostic tests are used to determine the infection

status of sampled animals and hence infer the infection status of the whole region. Unless all animals are sampled with a perfect diagnostic test there will be uncertainty in the subsequent estimate of the region's infection status. Much work has been done to determine optimal sample sizes and performance limits of such surveys [Cannon, 2002; Branscum et al., 2006; Cameron and Baldock, 1998a,b; Humphry et al., 2004; Johnson et al., 2004]. All these studies assume a probability based sampling of animal holdings, usually, within a defined *surveillance zone* which would normally surround the previously infected premises. Current disease control and surveillance policies generally involve the definition of a surveillance zone as a buffer zone a defined radius around previously detected, infected premises. In the case of foot-and-mouth disease in the United Kingdom this is a 10 Km radius zone around any previously infected premises. For the purposes of illustration and discussion the remainder of this chapter will focus on the design of post-foot-and-mouth disease epidemic surveys although the approach may be generalised to other diseases.

The World Organisation for Animal Health (OIE) have previously set prescriptive guidelines regarding the sampling strategy to be adopted for post-epidemic demonstration of freedom from foot-and-mouth disease. Recent changes in OIE guidelines introduce considerable freedom in how disease freedom surveys may be designed, allowing a more pragmatic approach to design, provided that the survey adequately supports a claim of disease freedom [OIE, 2009b].

Traditionally, farms are sampled from within the surveillance zone on a random basis to achieve an expected survey system performance (e.g. to have a 95% confidence of detecting an infected farm if infection is present at some predetermined, design value, such as 2%). Within each selected farm, samples will be taken from animals to achieve a within-farm expected survey performance - typically to detect infection on the farm with a 95% probability if it were present at some previously defined, within-group design prevalence (often 5%). More recently, in the case of disease control by vaccination, current procedures may require testing of all vaccinated animals on all vaccinated farms [TAIEX, 2007]. This more rigorous testing is aimed to overcome the possibility that vaccinated animals may have subtle clinical signs if infected and may have a higher likelihood of becoming carrier animals [Arnold et al., 2007; Paton et al., 2006].

This chapter examines the traditional approach for post epidemic surveillance using random selection of farms within the surveillance zone and suggests a more efficient methodology. This involves sample selection from wider areas using a latent infection risk model to choose which farms to sample and when to sample them. This approach is expected to reduce sampling costs by requiring fewer farms to be tested and to expedite the return to disease-free status by finding infected farms, if they are present, more quickly.

2.3 Methodology

The mathematical nomenclature and symbols used in this chapter are listed in table 2.1.

Variable	Symbol
A positive test result for farm/region	T^+
A diseased farm/region	D^+
Number of farms in restricted zone	N
Probability of farm i being latently infected on day t	$p_{(i,t)}$
Probability of farm i being latently infected ever	P_i
Farm level sensitivity (farm i)	ρ_i
System sensitivity	SSe
Probability that a sampled farm ultimately tests positive	τ
Sampling cost of ith farm	c_i
Sample size	n
Indices of n sampled farms	$s_{1..n}$
Sampling cost budget	C

Table 2.1: Nomenclature used for terms in this chapter

2.3.1 Theory

2.3.1.1 Surveillance system sensitivity

The purpose of post-epidemic sero-surveillance is to inform a decision about the disease status of an area under consideration, normally the restricted area. Return to international trade and removal of disease control measures are likely to be contingent on this decision. The area will be considered free of disease if it is contains no infected animals or evidence of circulating infection. Hence the performance of a survey to demonstrate disease freedom can be assessed by its ability to detect infection in an area if present. Martin et al. [2007b] call this the *system sensitivity* (SSe). For this analysis I assume that an optimally efficient survey design either:

1. Maximises the probability of detecting any previously undetected infected animals in the area (SSe), if present, for a given budget or,
2. Minimises the sampling cost of a survey that achieves a desired probability of detecting previously undetected infected animalsSSe, if present, in the area

The system sensitivity SSe is defined:

$$SSe = Prob(T^+|D^+)$$

Where $Prob(T^+)$ is the probability of at least one farm testing positive in the region, $Prob(D^+)$ is the probability of at least one farm being infected. $1 - SSe$ is effectively the probability that a surveillance system will fail to detect infection in an infected region.

I assume perfect system specificity as discussed by Martin et al. [2007a]. Any initially test-positive farms are retested until they are demonstrated to be genuine positive or cannot be shown to be negative. Effectively the case definition for a positive farm is a farm that tests positive after confirmatory and follow-up retesting i.e.

$$Prob(T^-|D^-) = 1$$

In order to estimate SSe:

$$Prob(T^+) = Prob(T^+|D^+) \times Prob(D^+) + Prob(T^+|D^-) \times Prob(D^-)$$

but since

$$Prob(T^+|D^-) = 1 - Prob(T^-|D^-) = 0$$

$$Prob(T^+) = Prob(T^+|D^+) \times Prob(D^+)$$

And hence

$$SSe = \frac{Prob(T^+)}{Prob(D^+)} \tag{2.1}$$

If the probability that each farm is latently infected (i.e. was infected but the infection is not detected during the epidemic) can be estimated, the system sensitivity (SSe) of different post epidemic sampling strategies can then be calculated using 2.1. It is assumed that any previously detected and infected farms have been removed from the sampling frame. If P_i is the probability that farm i is latently infected, $s_{1..n}$ are the indices of a sampled set of n farms from the population of N farms and ρ_i is the farm level sensitivity on farm i (i.e. the probability of a positive farm result if an infected farm is selected, sampled and tested). For this analysis I have assumed that P_i and ρ_i are independent, i.e. the sensitivity with which infection is detected on a farm is independent of the risk that the farm is latently infected. This assumption is discussed in section 2.5.8.

$$\begin{aligned} SSe &= \frac{Prob(T^+)}{Prob(D^+)} \\ &= \frac{1 - \prod_{i \in s_{1..n}} (1 - P_i \times \rho_i)}{1 - \prod_{i \in 1..N} (1 - P_i)} \end{aligned} \tag{2.2}$$

2.3.1.2 Risk based versus random farm selection

If all farms have an equal probability of being latently infected, the designed survey has a constant farm level sensitivity across all farms and the cost of sampling farms is constant there will be no advantage in selecting farms for sampling. However if latency probability P_i or farm sensitivity ρ_i vary then maximum system sensitivity within a sampling or cost constraint will be obtained by selecting a sample set $s_{i=1..n}$ of n farms to maximise $Prob(T^+)$:

$$1 - \prod_{s_{i=1..n}} (1 - \rho_i \times P_i)$$

As $\rho_i \times P_i$ is always positive this is maximised by maximising $\prod (\rho_i \times P_i)$ given the constraint of a number of farms to be sampled.

If the sampling constraint is simply a number of farms n to be sampled this gives a sample set of the first n farms when they are ordered by decreasing $(\rho_i \times P_i)$.

If the sampling constraint is a fixed sampling cost budget C and it is assumed that all farms have an equal sampling cost c^* (cost of visiting the farm plus cost of sampling on the farm) then the most efficient sample will be the first n farms ordered by increasing $\{1 - \rho_i \times P_i\}$ where n is at a maximum but satisfies $nc^* < C$.

2.3.1.3 Efficiency gains from risk based sampling

For risk based sampling to give an increased system performance (with a sampling/cost constraint) there needs to be both a variation in $P_i \times \rho_i$ and a predictive model that allows a survey design to select farms with a comparatively high $P_i \times \rho_i$ compared to random selection. In the following section I adapt a previously published model for risk of infection and use it to estimate potential gains in surveillance system efficiency resulting from application of risk-based as opposed to random sampling in a post-epidemic surveillance scenario.

2.3.2 Application to post-epidemic surveillance

2.3.2.1 Study scenario - Devon, UK 2001 foot-and-mouth disease epidemic

The foot-and-mouth disease outbreak in Devon 2001 was estimated to start on the 17$^{\text{th}}$ February and end with animals on the last known infected premises culled on 19$^{\text{th}}$ June 2001. Sero-surveillance during and after the epidemic identified 15 farms in this area as previously infected (recorded in the DEFRA submission to the OIE to substantiate freedom from disease and recorded in the disease control system database [DEFRA, 2002] and [Gibbens and Wilesmith, 2002]).

I adapt a model of the UK 2001 foot-and-mouth disease epidemic from Keeling et al. [2001] using parameter estimates from Diggle [2006] to estimate the probability of infection for each animal holding in Devon. From the model the probability that an uninfected farm i becomes infected on day t of the epidemic is estimated using the formula:

$$p_{i,t} = 1 - exp\left(- \sum_{j \in \text{infected farms at time } t} \lambda(t)\eta(d_{ij})A_j B_i\right)$$

Where:

$$A_j = 1.42N_{cattle}^{0.13} + N_{sheep}^{0.13}$$

$$B_i = 36.2N_{cattle}^{0.13} + N_{sheep}^{0.13}$$

A_j represents the transmission potential of an infected farm j, B_i the susceptibility of farm i. N_{cattle} and N_{sheep} are the stocking numbers of cattle and sheep respectively. The parameter $\lambda(t)$ is a baseline hazard.

The distance kernel $\eta(d)$, representing the decreased risk of transmission between farms with increasing straight-line separation, d_{ij} is modelled:

$$\eta(d) = exp\left\{-\left(\frac{d}{0.41}\right)^{0.5}\right\} + 1.3 \times 10^{-4}$$

This gives a declining relative probability of transmission with distance as shown in figure 2.1

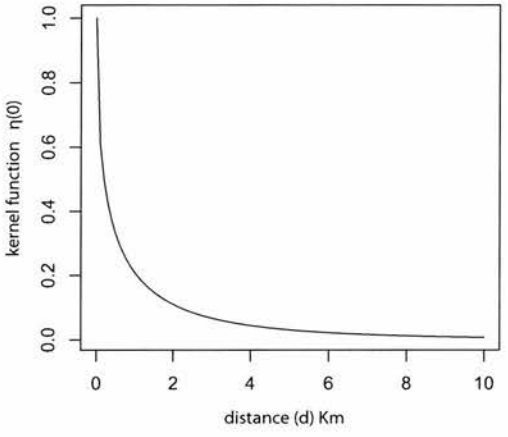


Figure 2.1: The distance Kernel used in the model [Diggle, 2006]. d is the distance separating an infectious and a susceptible farm, $\eta(d)$ is the relative probability of transmission.

2.3.2.2 Model parameters

Model parameters (see table 2.2) were as used in [Diggle, 2006]. The baseline hazard ($\lambda(t)$) represented the overall risk of transmission occurring in the Devon area, varying with time, during the epidemic. In Diggle [2006]'s study this was estimated from the daily case report data using a survival analysis technique (approximately 4.88 e-05 with slight variation over the course of the epidemic). This baseline hazard would be dependent on stocking densities, husbandry practices, foot-and-mouth disease virus serotype, time of year and other factors. For the Devon region the baseline hazard was relatively constant with time so was kept fixed in my model for simulation purposes. To assess the robustness of my estimates of risk based surveillance the baseline hazard parameter lambda was varied between simulation sets from 1×10^{-5} to 5×10^{-4}.

Variable	Value
Relative susceptibility in cattle	36.2
Relative susceptibility in sheep	1
Relative transmissibility in cattle	1.42
Relative transmissibility in sheep	1
Power correction	0.13
$\lambda(t)$	4.88 e-05 (base value)

Table 2.2: Parameters for model to predict infection probability

The modified Diggle [2006] model is used to estimate the probability of infection in each of the 4856 premises with sheep recorded in Devon in the 2001 UK agricultural census [DEFRA, 2000] using the 172 infected premises from the DCS database [Gibbens and Wilesmith, 2002] as sources of infection. The results from this model are shown in section 2.4.1

In the study I need to predict the probability that farm i will become infected and that infection will *not* be clinically discovered (or serologically discovered during routine sero-surveillance prior to formal post epidemic sero-surveillance). There are two scenarios by which an infected farm will become a latent (undiscovered) infected farm by failure to observe clinical disease: one is if infection on the farm is not clinically detectable i.e. disease is effectively sub-clinical; the other is if clinical signs are present but are not observed. The latter scenario is partially time dependent, there being a limited period during which infection may be clinically detected (of up to several weeks duration). As serological surveillance will normally start several weeks after the last clinically observed case and the frequency of cases in the tail end of the epidemic is low the time dependent effects may be disregarded and the two modes by which a farm may be missed considered as one probability. Hence in order to use the model [Diggle, 2006] it is necessary to estimate a *non-discovery* multiplier to allow for the probability that an infected farm is not clinically detected during the first few weeks post-infection and then goes on to become a latently infected farm. This probability may be different for every farm and is likely to change over the course of the epidemic and post-epidemic period as surveillance efforts vary. Unfortunately only limited data are available to estimate this parameter for each farm.

2.3.2.3 Estimation of probability of discovery of infection

Sero-surveillance in Devon, after the 2001 epidemic, involved sampling and testing of some 4,407 farms out of the estimated 5,000 remaining sheep farms [DEFRA, 2002]. For the purposes of this analysis I assume that all farms that could be serologically discovered were discovered. The sero-surveillance identified ten holdings as seropositive in Devon after the slaughter date of the last infected premises (I have not considered the five serologically detected premises prior to this date as representative of latent infections). 172 holdings were identified as infected premises during the epidemic in Devon having been detected either clinically or by in-epidemic sero-surveillance. Exploratory data analysis suggests that farms that were discovered during the epidemic were more likely to have cattle present and more likely to be close to other infected premises. A logistic regression model estimated in a Bayesian framework was used to model the heterogeneity of discovery probability across farms whilst capturing the uncertainty in these parameter estimates resulting from the limited data. Given the limited number of serologically discovered farms in the data set I elected to use binary predictor variables of cattle presence (i.e. cattle numbers greater than zero) and adjacency to the nearest IP (less than 3 Km) to reduce the model parameter count. The 3 Km adjacency cut-off was selected as regulations require there to be a heightened surveillance during an epidemic in the area within 3 Km of infected premises and hence likely to be a higher probability that an infected farm is discovered. The probability of detection is assumed to remain fixed for each farm over the course of the epidemic in the absence of sufficient data to estimate a time dependent effect.

The probability of discovery of an infected farm was modelled:

$$Prob\,(\text{discovery}) \;=\; \frac{e^z}{1+e^z}$$

$$z \;=\; \beta_0 + \beta_{cattle} X_{cattle} + \beta_{near} X_{near}$$

Where $exp(\beta_0)$ is the baseline odds of discovery, $exp(\beta_{cattle})$ the odds ratio for discovery if cattle are present and $exp(\beta_{near})$ the odds ratio for discovery if the farm is within 3 Km of an IP. X_{cattle} and X_{near} are indicator variables for presence of cattle and adjacency to an IP.

The parameters of the logistic regression model were estimated from the sero-surveillance results for Devon in 2001 using the JAGS software [Plummer, 2003] to produce 5000 sets of samples from the posterior distribution of the parameters after a burn-in of 5000 simulations. Alternative parameterisation and the inclusion of different predictors were compared using the Deviance information Criterion [MRC Cambridge]. Chain convergence was assessed using the Brooks, Gelman and Rubin statistic [Brooks and Gelman, 1998].

Draws from these posterior distributions were then used with demographic data from the 2001 Agricultural census (as a sampling frame and source of stock/location data) to estimate the mean probability of discovery, if infected, for each sheep farm in Devon.

The discovery model does not account for farms that were infected, not discovered during the epidemic but subsequently culled as part of the disease control process (where farms contiguous or otherwise associated with infected farms where pre-emptively culled). If a future epidemic used a similar control process the results from this study would be valid as such undiscovered farms would face a similar risk of being culled. Different future disease control processes would bias the performance estimates based on this study.

Probability of infection from a published model [Diggle, 2006] and probability of discovery if infected from the above model are combined to estimate the *daily* probability for each sheep farm that the farm has become infected and that the infection has not been detected. This assumes that the probability of discovery and probability of infection are independent, conditional on the risk factors used in the infection and discovery models.

The resulting *daily* probability of farm i becoming infected yet not clinically detected is converted to a probability that farm i has become infected yet not clinically detected at *any* point during the epidemic, P_i, assuming $P_i(t)$ is independent (see section 2.5.5) using:

$$P_i = 1 - \prod_{t \in (\mathrm{T})} (1 - p_{i,t})$$

Where $p_{i,t}$ if the probability of latent infection on farm i on day t from the Diggle [2006] and discovery models and T duration of the epidemic.

2.3.3 System sensitivity versus number of farms sampled

The system sensitivity (SSe) of risk based surveys is estimated using a range of farm sample sizes (n) from just one farm to the whole population, by preferentially selecting the n farms with the highest probability of latent infection. This result assumes that the within-farm sensitivity is 100% i.e. if a latently infected farm is selected and sampled that infection, if present, will be detected with certainty. The results are shown in section 2.4.3.

2.3.4 System sensitivity of random sampling

For comparison the expected system sensitivity for randomly selected samples is estimated by simulation. The sampling is assumed to take place from within the 3 Km protection and 10 Km surveillance zones. For each sample size (n) of farms from 1 to the remaining, post-epidemic, population of 3526 sheep farms, within the protection and surveillance zones, 1000 random samples of size n were drawn (without replacement). The system sensitivity of each sample was calculated using equation 2.2 and the mean for each sample size stored. Additionally the worst case results are estimated by simulation of selection of farms in order of increasing risk of latent infection. This gives a lower bound of system sensitivity for a given farm sample size. The results are shown in section 2.4.3.

2.3.5 Zonal location of sampled farms

Conventionally farms are sampled from within the surveillance and protection zones. This may not be the most efficient approach; it is possible that farms outside these zones may be at higher risk of latent infection and hence should be sampled with priority. To investigate this I classify farms by decreasing probability of latent infection according to their location within their zones. The zones used are the current 3 Km protection and 10 Km surveillance zones constructed using the respective buffers around the 172 source infected premises from the Devon 2001 epidemic. The results are shown in section 2.4.4.

2.3.6 Consequences of imperfect information / robustness analysis

The above methodology makes the strong assumption that the model used to estimate each farm's probability of latent infection is correct. It does not assume perfect information about each farm's latent infection status but perfect information about each farm's *probability* of latent infection. In reality any model estimating the probability of infection in a farm will have error. In this section I investigate the sensitivity of the performance gains of risk based sampling to uncertainty in the performance of the predictive model.

With risk based sampling it is the rank of estimated probabilities and their heterogeneity that determines the choice and expected benefits of farm selection. A model that estimates the probabilities as a monotonic increasing function of the true probabilities will still be able to perfectly inform selection of an optimal sampling set (though give an incorrect estimate of its system sensitivity) for a given sample size. Sub-optimal farm selection will only occur when the model causes incorrect selection of lower probability farms for inclusion in the sample. Imperfect models that correctly rank the farms' risk of undiscovered infection will still, however, incorrectly estimate the system sensitivity achieved by a survey.

To investigate the effect of model uncertainty I have used two sets of models which have an error component. One modifies the estimated probability of infection for each farm using a transformed normal error component. The other set randomises a proportion of the estimated probabilities so that only a proportion of the farms' have an accurately estimated risk of infection. These models are then used to select farms and with the original, error-free model to estimate the system sensitivity performance of the resulting sample sets:

2.3.6.1 Transformed normal error approach

The Diggle [2006] model and discovery model are used to estimate the probability of latent infection Pi for each farm i. This value is then transformed to the log-odds scale and a normally distributed error component is added. The result is then re-transformed to the probability scale:

$$P_i^t = invLogit\left(logit(P_i) + N(0, \sigma)\right)$$

Where P_i^t is the transformed probability of latent infection on farm i, *logit* is the logistic transformation $logit(x) = log(1/(1-x))$ and *invLogit* the inverse $invLogit(x) = 1/(1 + exp(-x))$. For σ greater than zero this will add an error component to the model's predictions. For a range of σ from 0 to 25 the estimated probabilities with error were used to select farms for sampling and the probabilities from the original [Diggle, 2006] model were used to estimate the resulting system sensitivity (performance). This estimation was repeated 1000 times to provide an estimate of the mean performance with each level of model error. The results are described in 2.4.5.

2.3.6.2 An alternative partial knowledge approach

Rather than simulating errors in the estimation of each farm's risk I also estimate the performance of risk based sampling if I can only correctly estimate the risk for a proportion of all farms. This was done for proportions from 1 (i.e. perfect knowledge) to 0 (i.e. no information on risk - random selection). The results are shown in section 2.4.6.

2.3.7 Effect of risk based sampling on delay to declaration of disease freedom

Given a selected set of farms for sero-surveillance sampling (whether by the above model-driven method or by another, e.g. random or a veterinary expert directed method) there may be flexibility to choose the order with which farms are visited and sampled. Sampling and subsequent sample handling, analysis and recording may introduce delays of several days between sampling and result. If a sampled farm tests positive for infection and virus is subsequently isolated on a farm it may be classified as a new outbreak and consequently disease freedom may not be declared until a fixed period has elapsed after the culling of stock on this farm. Hence it is desirable to order sampling such that farms that are most likely to be latently infected will be visited, sampled and analysed first.

To estimate the potential benefits of *ordered* sampling I estimate time from start of sampling to identification (on average) of the last positive farm for sero-surveillance in Devon. A sample set of farms of a given size are selected either at random from the whole population or by decreasing risk of latent infection using the model. sampling from these farms is then simulated over a surveillance period assuming that 50 farms can be sampled each day. In the risk based approach high risk farms are sampled first. In the random approach the farms are sampled in a random order. Then over repeated simulations the status of sampled farms is simulated (using a bernoulli process with probability equal to the farm's probability of latent infection). For all farms this status is combined with their simulated sample timing data to give the time that each positive farm was sampled. Hence the maximum of this allows estimation, for each sampling approach, of the time of sampling of the *last* positive farm. Repeated over the simulation set this gives an estimate of the time from start of sampling to the sampling of the last positive farm for each approach (d).

$$d = \max \{\text{sampleDate}_i \times bern(P_i)\}_{i \in \text{farms to sample}}$$

Where d is the date of sampling of last positive farm, sampleDate_i the sampling date for farm i and $bern(P_i)$ a bernoulli random variable with probability P_i. I assume that 50 farms are sampled and tested each day.

The simulation was repeated 5000 times for a range of sample sizes and the mean of the latest dates stored. The results are described in section 2.4.7.

2.4 Results

2.4.1 Results from latent infection model

The model, with the base baseline-risk parameter ($\lambda_0 = 4.88 \times 10^{-5}$), estimated an expected 11.2 latent farms in Devon with 7.8 within the protection zone (0 – 3 Km from the nearest IP), 3.0 within the surveillance zone (3 – 10 Km from the nearest IP) zones and an expected 0.4 farms outside these zones. The spatial distribution of risk of latent infection is shown in Figure 2.2. Farms with a high probability of harbouring a previously undiscovered infection are spatially associated with the premises that were

identified as infected during the epidemic. The spatial component of risk of latent infection will be a combination of the increased risk of farms *near* to detected infected premises and the increased probability of undiscovered farms being *further* from detected infected premises. The results suggest that the modelled risk overwhelms the discovery effect such that overall farms near to detected infected farms are more likely to be latently infected.

Table 2.3 gives the mean result of 500 simulations for the 3 km protection zone, the 10 km surveillance zone and the whole region. Results are given for each zone and cumulatively across zones. They are the number of farms, the expected number of latently infected farms and the system sensitivity if all farms in the zone were sampled with an on-farm survey of 100% sensitivity.

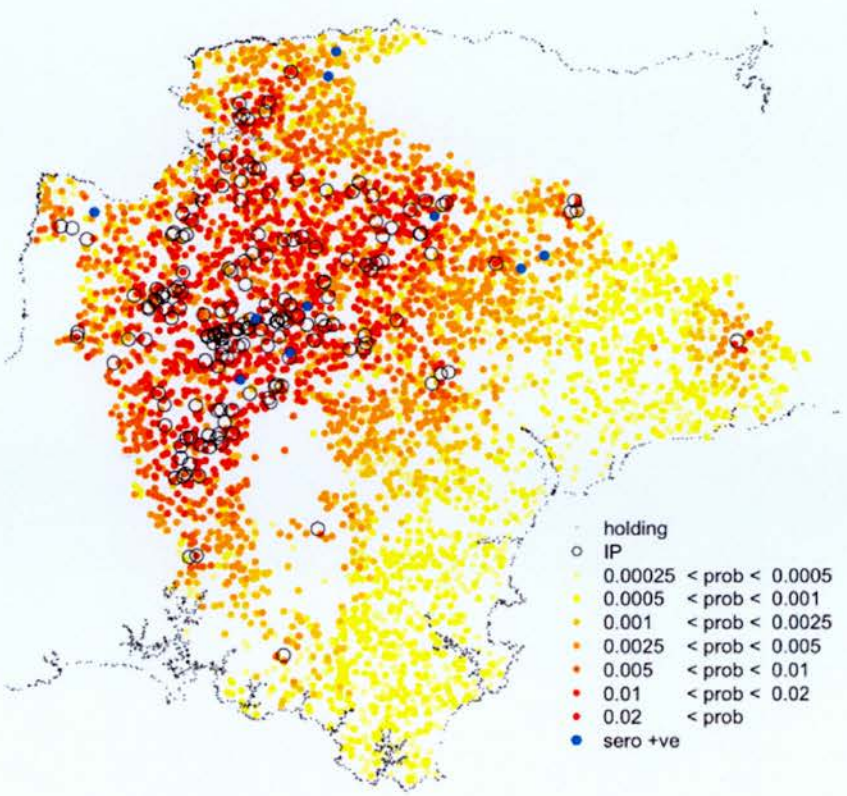


Figure 2.2: Estimated risk that individual sheep/mixed farms may be latently infected with foot-and-mouth disease at the end of the epidemic. The ten farms found to be sero-positive in 2001 after the epidemic are also shown (blue dots)

2.4.2 Results from discovery model

The results of the latent infection discovery model are shown in table 2.4. Farms with cattle present were much more likely to be detected, if infected, during the epidemic as were farms within 3 kilometres of a previously infected farm. The magnitude of these estimates has a large uncertainty.

2.4.3 Performance of risk based sampling

The comparative results of risk based sampling are shown in Figure 2.3. Risk based sampling gives a markedly better performance than random sampling for the same sample size (number of farms visited); for example, assuming the model provides perfect information about a farm's probability of latent infection, only 225 farms need to be sampled to give a 95% system sensitivity as compared to 1083 farms if random sampling is used. The system sensitivity performance gains relative to sample size would either increase the probability that the survey system detects disease, if present, or reduce the cost of surveillance to attain a required (95%) system performance. The gains of employing risk based sampling decrease as the sample size increases until approximately 2000 farms are sampled when both risk based and random sampling converge onto a system sensitivity of 1. Figure 2.3 also shows the worst case for sample design where farms were sampled in order of increasing probability of detecting latent infection, i.e. the lowest risk farms are selected with priority. This identifies a lower limit for the performance of the surveillance system, i.e. a worst case if farm selection was based on a model that ordered the risk of infection in farms as the complete inverse of the true risk.

Zone	Total farms	Expected latent farms	SSe (all farms sampled)	Cumulative Farms
Protection zone	1439	7.79	0.976	1439
Surveillance zone	2087	3.01	0.821	3526
Other	1330	0.381	0.337	4856

Table 2.3: Sheep farm distribution and model predictions

	Mean	SD	Odds
Cattle present	3.11	0.97	22.5
Near IP ($\leq$3 Km)	2.67	0.87	14.4

Table 2.4: Results from Bayesian discovery model - mean and SD of coefficients of logistic model (and odds ratio) predicting discovery of an infected farm.

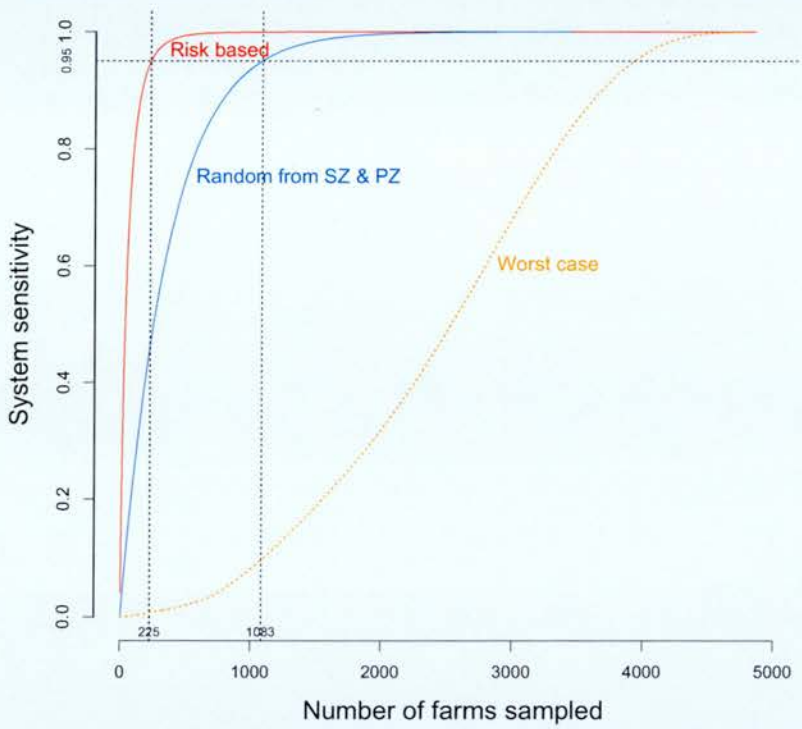


Figure 2.3: Comparison of system sensitivity of random (blue line) and risk based (red line) sampling. Horizontal dashed line is at 95% system sensitivity (the conventional target for farm level post-epidemic surveillance) with vertical lines showing corresponding sample sizes. (Worst case performance shown with an orange dotted line)

2.4.4 Zonal sampling

Conventionally, sampling has drawn samples, unless veterinary judgement suggests otherwise, from the 10 Km surveillance zone. Figure 2.4 shows the zone that samples would be taken from when risk based sampling is used with the potential to draw from any farm in the restricted area. As sample size increases to 1476 farms the first farm outside the protection and surveillance zones will be selected. By a sample size of 2254

1% of farms will be selected from the area outside the surveillance zone. These sample sizes represent system sensitivities of virtually 100% and as such would be unlikely to be required for regulatory purposes. Hence this suggests, given the risk model used in this study, that the conventional approach of sampling within the surveillance zone is rational if random selection within a geographical zone were a regulatory requirement.

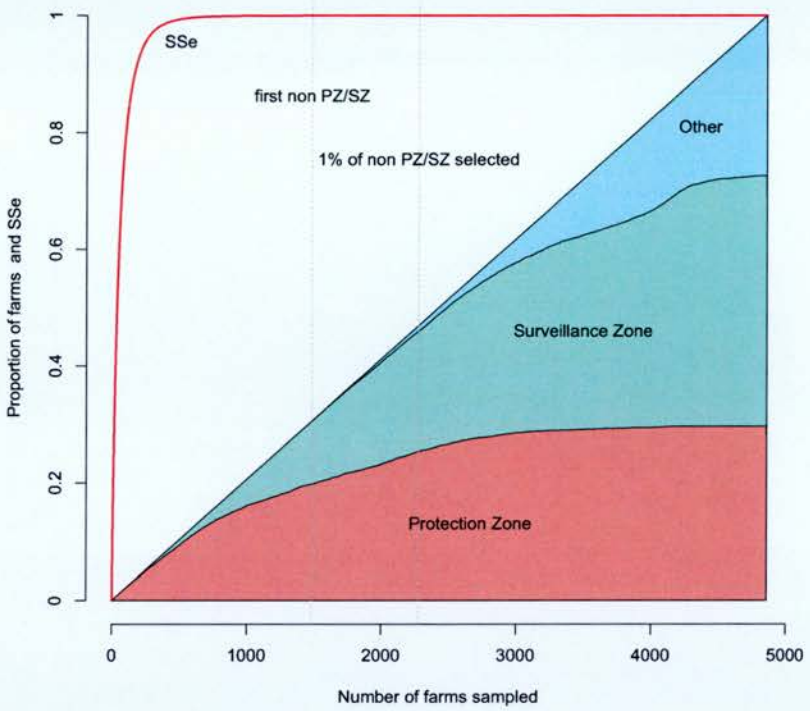


Figure 2.4: Source zone of farms when the farms are allocated in decreasing probability of latent infection. Vertical dotted lines showing points at which first and 1% of farms outside surveillance zone and protection zone will be selected. The red line shows the estimates system sensitivity (SSe) using risk based selection for the range of sample sizes.

2.4.5 Consequences of imperfect information - logistic error model

Imperfect models of farm risk will give a reduced performance for risk based selection for surveillance. In this section error was added to the model's estimated probability that each farm is latently infected. The error component was drawn from normal distributions with different variances. The error was added in the logistic domain. Figure

2.5 shows the effect of this error on the model estimates with increasing variance. The fitted regression lines show that, as expected, even with high added error the noisy data still retains the overall trend of the noise free estimates. Figure 2.6 shows estimated system performance loss of using risk based selection of farms with error introduced to the risk model. The performance versus sample size of risk-based selection is compared to the performance of random selection of farms from within the 10 Km surveillance zone. The uncertain models with high error will (when bad enough) perform worse, on average, when selecting from the whole region than random selection from protection/surveillance as these traditional zones do contain the majority of risk. A logistic noise standard deviation of 10 downgraded the risk based selection to have similar performance to random, surveillance zone based selection.

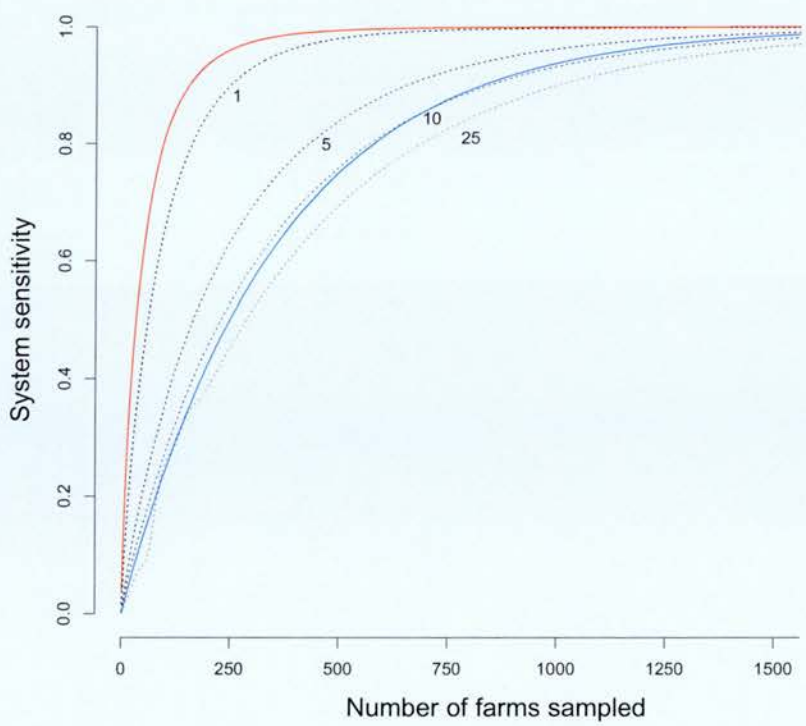
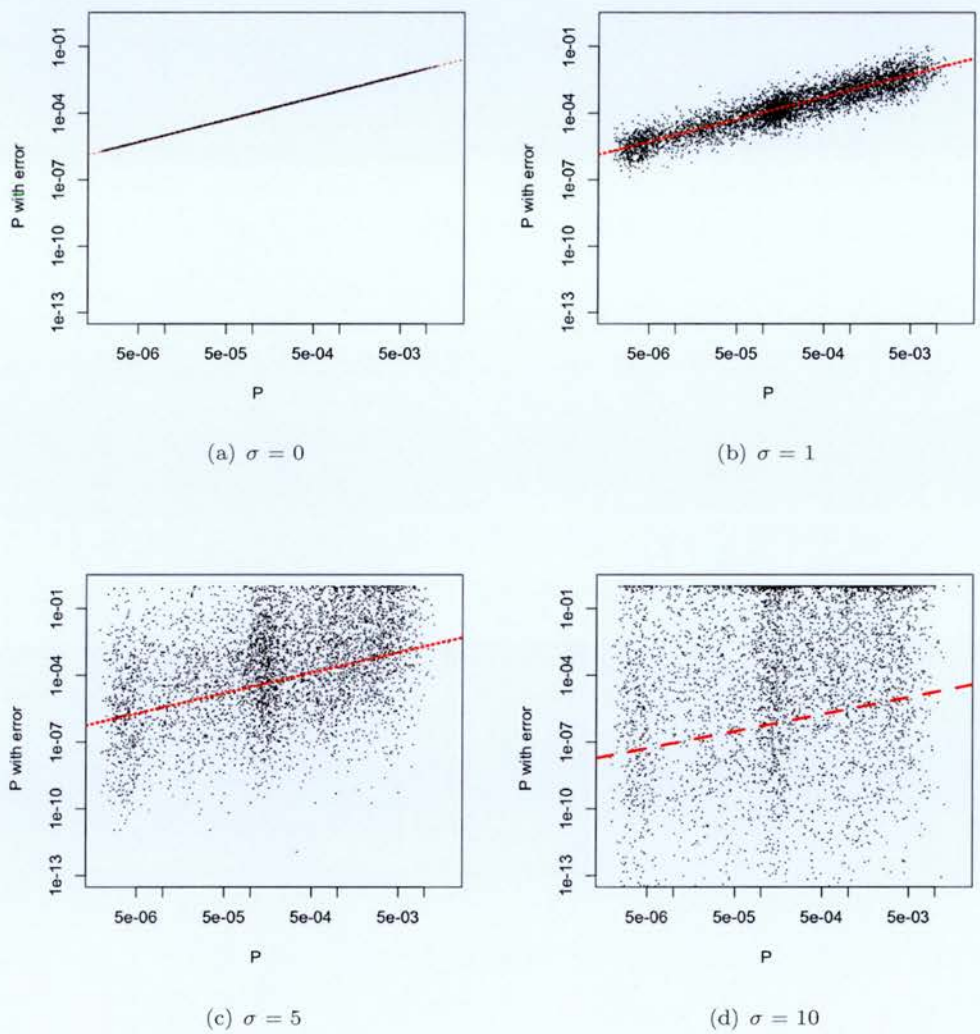


Figure 2.6: System sensitivity versus sample size (number of farms selected) for error-free, uncertain and worst case models versus random selection. Error-free model shown in red and blue line shows random selection from protection and surveillance zones) - uncertain model results shown in grey annotated with standard deviation of additive error component ($N(0, \sigma)$) on the log odds scale (σ from 1–25).



(a) $\sigma = 0$

(b) $\sigma = 1$

(c) $\sigma = 5$

(d) $\sigma = 10$

Figure 2.5: Probability of latent infection in farms in Devon (in ascending order - x-axis) plotted against this probability with an added logistic normal error component (P with error — P_i^t) ($invLogit(Logit(P_i) + N(0, \sigma))$) for different σ. The red dotted line is a fitted linear model (form $P_i^t = a + b \times P_i$).

2.4.6 Consequences of imperfect information - proportion known model

An alternative exploration of the impact of imperfect models is shown in figure 2.7. The method estimates the system sensitivity when only a given proportion of the farms have their risk known. Risk based sampling will perform more efficiently than random sampling from the 10 Km surveillance and 3 Km protection zones until the proportion of accurately known farm-risks of latent infection drops below approximately 20%.

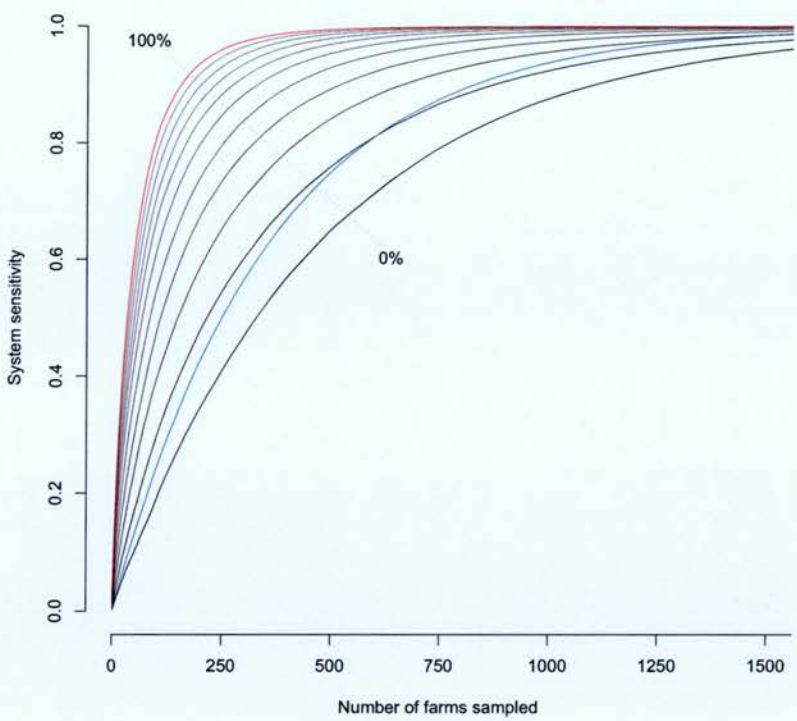


Figure 2.7: System sensitivity versus sample size (number of farms selected) for error-free and partial knowledge models versus random selection. Error-free risk based model shown in red. Blue line shows random selection from protection and surveillance zones) - partial knowledge model results, with selection from whole region, shown in grey labelled with percentage of farm where risk is known. The models range from 0% to 100% of farms' risk known in 10% increments.

2.4.7 Effect of ordering on delays to disease freedom

The time to last positive sample is plotted against the system sensitivity for a risk-ordered and randomly ordered approach to farm sampling in figure 2.8. The result demonstrates that risk based sampling and ordering, given a desired system sensitivity, will result in any farms that turn out to be test positive being selected and sampled markedly sooner than if ordering is not used. For a target system sensitivity of 95% risk based selection and ordering will, on average, mean that the last positive farm is sampled at 2.8 days after sampling commences whereas random selection will mean that the last positive farm is sampled at 18.4 days. Detection of a test positive farm will result in follow up confirmatory tests and possible disease control consequences such as further investigation of contact farms and animal culling. Hence risk based ordering may decrease delays to ultimate declaration of disease freedom in the region.

2.4.8 Sensitivity analysis with respect to baseline risk on infection

The previous results use a baseline hazard ($\lambda = 4.88 \times 10^{-5}$). Changes in the baseline hazard will shift the estimated latent probabilities for all farms in the area with consequent effects on the estimates of system sensitivity for different sample sizes. This effect is illustrated in figure 2.9 where the system sensitivity of risk based and random selection is shown against sample size for three different baseline hazards. As the baseline risk decreases the relative benefits of choosing farms with a high probability of latent infection increase. Or alternatively, as risk increases it becomes less important to focus surveillance — intuitively, infection is more likely to be present on any randomly selected farm so it matters less where you look.

Table 2.5 shows this quantitatively. As baseline hazard decreases there are fewer, on average, latent farms in the region. More samples are required to achieve a target system sensitivity and hence the relative benefits of risk based surveillance increase. Additionally the time saved in delay to sampling the last positive farm compared to random selection increases as baseline risk decreases; i.e. risk based selection and ordering becomes increasingly beneficial to both performance and timing as baseline hazard decreases.

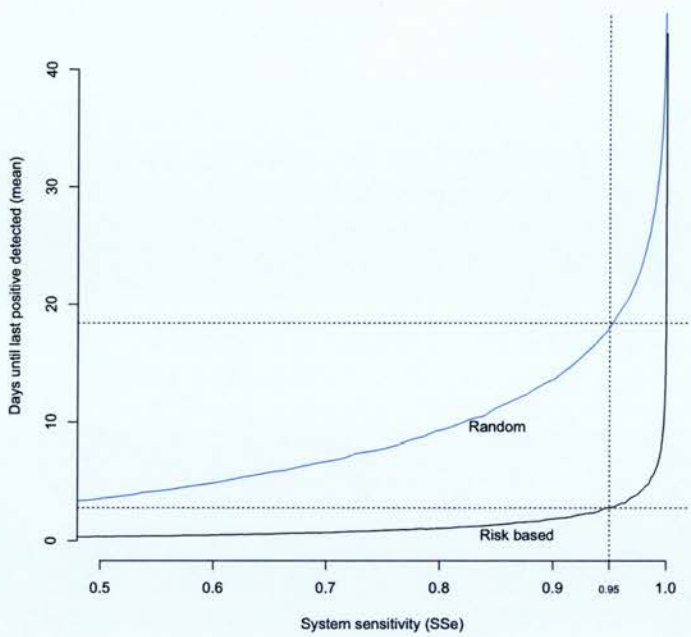
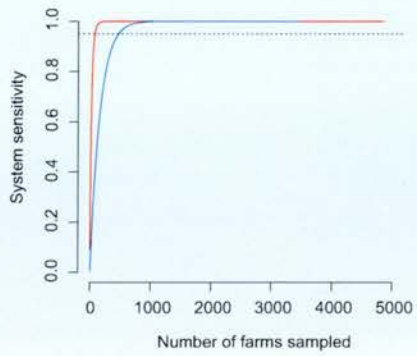
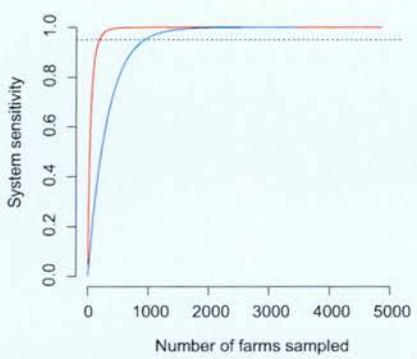


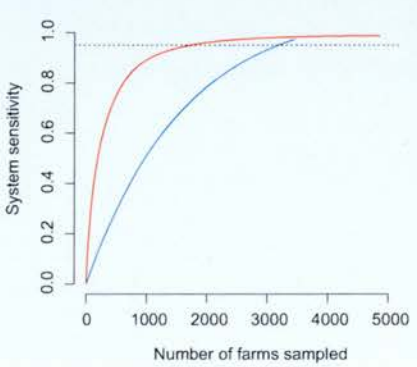
Figure 2.8: Expected time to last positive farm for randomly selected/ordered and risk model selected/ordered sampling.



(a) $\lambda = 10 \times 10^{-5}$



(b) $\lambda = 5 \times 10^{-5}$



(c) $\lambda = 1 \times 10^{-5}$

Figure 2.9: Effect of decreasing baseline risk of farm infection on the performance of surveys with increasing sample size. System sensitivity (SSe) of risk based selection is shown in red and random selection is shown in blue.

λ	Latent farms (in region)	Samples (Risk based)	Samples PZSZ (Random)	ΔDelay
5e-04	111.1	13	103	12.1
1e-04	24.0	75	478	22.6
5e-05	12.0	184	945	26.9
base	11.0	225	1083	27.0
1e-05	2.0	1934	3337	25.8

Table 2.5: Effect of baseline risk. The table shows the effect of decreasing baseline hazard (reference value was 4.88×10^-5) on the number of latent farms in the region, the number of samples required to achieve 95% system sensitivity for risk based and random selection (from the surveillance zone) and the reduction in delay (ΔDelay) to finding the last positive farm (if all farms are sampled using risk based compared to random ordering). The (ΔDelay) measures how targeted sampling will speed up detection of any positive farms — a higher number represents a greater benefit to risk targeted sampling.

2.5 Discussion

This study shows that a suitable model of infection risk can be used to target and order sero-surveillance sampling to increase efficiency and reduce delays to declaring disease freedom after an foot-and-mouth disease outbreak. Historically, post-epidemic surveillance for foot-and-mouth disease has targeted sampling to farms within the 3 kilometre protection zone and 10 kilometre surveillance zone, presumably on the basis that any undetected infected farms are most likely to be in these areas. This involves an implicit spatial assumption about disease transmission which may be derived from veterinary assessment. Risk based selection refines this to target specific farms with the vicinity of the outbreak/epidemic . The benefits of this technique are shown to be potentially large with an 80% reduction in sample sizes for the same system performance (95% SSe) and reduction in timing to last positive case of 15 days.

2.5.1 Model uncertainty

The realisation of these potential benefits is dependent on the accuracy with which a predictive risk model can order farms by infection risk. I have used simple approaches to model uncertainty to explore the consequences of imperfect knowledge of farm infection risk. This shows that the benefits of targeted sampling were reasonably robust to the introduction of noise/uncertainty into model estimates; knowledge of only 20% of farms' infection risk will provide a system of better performance than surveillance-zone targeted sampling. As one would intuitively expect, as long as a model has some information value it will increase efficiency and reduce delays to declaring disease freedom.

However, if a predictive model of farm infection risk is particularly poor its application could result in a survey that's efficiency and timing would be worse than a randomly sampled and ordered survey that draws from the traditional (overall high risk) surveillance and protection zones. Whilst such a scenario is unlikely it means that the risk based sampling and ordering approach should be used cautiously, especially with models that are potentially over-fitted, unstable or otherwise suspect.

Bias in the model estimating risk may result in expected gains of sampling to be under or over estimated as the quantitative benefits will depend on the degree of heterogeneity in the farm infection risk probabilities. As previously discussed, a *flat* risk distribution

will give no benefits to risk based surveillance whereas a risk distribution with a small proportion of very high risk farms will give great rewards to risk based surveillance. In the case of a biased risk model, whilst the risk-based technique may still propose an optimal set of farms for sampling the actual estimated benefits from my analysis are contingent on the distribution of risk of infection in both the region and the sample set.

The study assumes knowledge of the probability that infection is discovered on infected farms. In the analysis this uses data obtained from post-epidemic sero-surveillance. In real application these data would not be available. Available options are to assume that discovery probability is constant across all farms or to use historical data to estimate discovery probability under the assumption that this probability is similar for similar farming systems and diseases. These assumptions may reduce the estimated benefits of the targeted surveillance techniques.

2.5.2 Fixed cost assumption

The study approach assumes that the marginal cost of visiting a particular farm for sampling is constant. In reality the cost of sampling a particular farm will be conditional on which other farms have been sampled due to travelling costs and deployment of sampling teams. Hence it may be more efficient to use a sub-optimal sampling set and ordering if the consequent increase in required farm numbers and delays to freedom were outweighed by the reduced costs of sampling adjacent or convenient farms. This point does not negate the concepts raised in the study but suggests that in application it may be appropriate to extend the models to include sampling/travelling logistics. The cost of farm sampling is potentially more variable than latent infection risk depending on available resources, overall epidemic extent and geographical location.

Regarding the extension to our approach to include variable cost data: I have assumed that the unit cost of visiting a farm is equal for all farms. The optimisation is then one of benefit maximisation (in this case for system sensitivity) for a given cost budget. In reality sampling costs may vary from farm to farm and hence designing a survey that gives highest sensitivity for a fixed budget becomes more challenging. With a fixed cost, choosing farms in decreasing order of risk until a budget limit is exhausted will give the best expected performance for a budget.

When costs are allowed to vary it is necessary to maximise:

$$\frac{1 - \prod_{i=s_{1..n}} (1 - P_i \times \rho_i)}{1 - \prod_{i=1..N} (1 - P_i)} \tag{2.3}$$

for a given cost budget such that

$$\sum_{i=s_{1..n}} c_i \leq C$$

Where $P_i \times \rho_i$ is the product of risk of undiscovered infection and farm sensitivity on the ith farm, $s_{1..n}$ is a sample of n farms, N is the number of farms in the area and c_i is the cost of sampling /testing on the ith farm and C is the cost budget. Removing elements from 2.3 that are independent of the selected sample this means maximising:

$$\prod_{i=s_{1..n}} P_i \times \rho_i$$

given the cost constraint.

There is no simple rule to select farms for inclusion in a sample given their estimated risks (P_i) and costs (c_i). With several thousand farms as candidates for a survey which may be cost limited to roughly hundreds of farms there are too many combinations of possible sample sets to allow an exhaustive search. This search is further complicated as the costs of visiting a particular farm will be conditional on which other farms are being visited. An optimal sampling strategy if cost data were available would require carefully designed models and search algorithms to determine most cost-efficient sample sets.

2.5.3 Assumption of 95% farm sensitivity

For the graphical comparisons of performance I have assumed that all farms' livestock are sampled and tested to give a farm level diagnostic sensitivity of 95% in accordance with previous OIE codes [OIE, 2005b]. Depending on the size of the farm and the available diagnostic tests the cost of reaching this performance will vary from farm to farm and may even be unobtainable when the individual animal diagnostic test

sensitivity is lower than the farm level performance target and the farm size is small [Greiner and Dekker, 2005]. If this constraint is removed then a survey design would have freedom to choose not only which farms in the region to sample but rather which animals in the region to sample and indeed which diagnostic test to use and how to interpret the results for each sample. Although the removal of these constraints would give a potentially lower cost method of achieving a required regional performance it is a high dimension optimisation problem that is computationally difficult. Furthermore, the resulting varying farm by farm performance and sampling requirements may be politically harder to justify to both farmers and decision makers.

2.5.4 Spatial correlation — disease clustering

This analysis has used a spatial model to estimate the probability that farms are latently infected after an epidemic. However the interpretation of surveillance results disregards any spatial correlations between farms' risks. Foot-and-mouth disease infection at a farm level tends to be spatially clustered, i.e. a farm near an infected farm is more likely to be infected than a random farm. Thus by sampling and testing a farm more information is gained about the infection status of nearby farms than about distant farms. A risk based approach will tend to focus surveillance around high risk areas, such as previous infected premises. Given two farms of equal risk of latent infection more information would be gained, if disease is spatially clustered, by sampling a farm distant from a previous sample site than the one near to it. In the Devon scenario the infected premises that were identified during the epidemic were relatively dispersed and the risk based sampling model generates a dispersed sampling field. Hence estimates of performance benefits in this setting are likely to be reasonable estimates of true gains. A strategy to explore the issue could consider the explicit spatial correlation of risk in sampled farms. However, as the spatial correlation information of a farm to add to the sample set is contingent on the existing members of the sample set it would be necessary to use an iterative approach to pick sample sets, testing sets which do not necessarily contain the highest risks of latent infection but that may give more information through spatial correlation. Such an analysis would be numerically complex, potentially taking a factorial, $N!$, computational steps where N is the desired sample size. These problems can be computationally intractable [Harel, 1987].

2.5.5 Assumption that the non-discovery probability residual error is conditionally independent of residual in disease risk model

To estimate the probability that a farm is infected but has not been detected I have used a model that estimated probability of infection and detection and a model of non-discovery probability. The calculation of probability of undiscovered infection using the simple product of probabilities from these two models assumes that the residuals or error in these two models are independent. If the errors are correlated the final risk estimate will be biased.

2.5.6 Combination of multiple risk models

As already mentioned the risk-based strategy is, in effect, a development of the existing strategy that focuses surveillance on farms within the surveillance and protection zones. The existing strategy implies a 10 Km radially based model of infection risk. It is possible that there are other models of infection risk, which may be mathematical, empirical or conceptual, that one may wish to employ to inform sampling. These may be combined to design a sampling strategy by using appropriately weighted combination of the models' estimates to give a synthesis of their risks and using this to select and order sampling.

2.5.7 Effect of testing and control during the epidemic

To simplify the analysis I have not attempted to model the ongoing culling and testing that would normally occur during an outbreak; none of the at-risk population are considered to have been declared as infected premises during the outbreak. In a real outbreak farms would be culled, neighbouring farms would be subject to ongoing clinical and serological surveillance and latent disease may thus be discovered before the post-epidemic surveillance period. Nonetheless, although these discoveries may alter the estimates of the benefits of risk based and ordered sampling the risk based sampling approach will remain more efficient and bring an, on average, faster return to declaration of disease freedom than randomised sampling. However the estimates of benefits from the current study may over or under estimate the benefits under different control strategies.

2.5.8 Assumption that farm level sensitivity and probability of latent infection are independent

Risk of infection and probability of discovery of infection during an epidemic were estimated in this analysis on a farm by farm basis. However the sensitivity of on farm detection, given a farm was selected for investigation was fixed at 95%. This is the target on farm value based on previous OIE recommendations [OIE, 2005b] that individual animals be sampled within each selected epidemiological unit to give a 95% probability that an infected unit would be detected if 5% of animals on that unit were infected. Even if each farm survey is designed to meet this criteria there will be farm by farm variation in the true farm level sensitivity. Probabilities of latent infection on a farm by farm basis varied (as estimated by the model) from approximately 2×10^{-6} to 2×10^{-2} whereas on farm sensitivity is unlikely to vary beyond 20–100%. For example, a sensitive test applied to all animals in a group will give a farm level sensitivity approaching 100%, a poor (50% sensitive test applied to a minimum sample size of 40 animals out of a group of 100 with only one animal infected would still give a farm level sensitivity of 20%). So variation in latent risk probabilities is likely to overwhelm variations in on farm sensitivity. Hence any dependence of latent probability and farm level sensitivity will not affect the first order magnitude of the estimates of performance gains.

2.5.9 Concluding remarks

Current, random sample based surveillance is likely to be inefficient requiring more farms to be visited and more animals to be sampled than necessary to achieve a given performance target. Risk based selection of farms for post epidemic surveillance is more efficient and will also expedite the process of declaration of disease freedom. Prompt detection of latently infected farms may also reduce the risk of onwards transmission and recrudescence of the outbreak. Whilst risk based selection is robust, to give the best and most reliable gains, risk models also need to be robust and precise hence continued research will be required to model the disease and surveillance processes to best inform surveillance.

Chapter 3

Thrace sero-surveillance: Introduction and survey performance

3.1 Summary

The next three chapters examine the application of serology based surveillance for foot-and-mouth disease in the Thrace region of Turkey. Thrace is an important region as it lies between the Asian region of Turkey, subject to repeated foot-and-mouth disease incursions and Europe that remains generally free from foot-and-mouth disease without vaccination [OIE-WAHID, 2008]. A foot-and-mouth disease vaccination program in Thrace is currently in use to provide a buffer for disease control. Annual or biannual surveys are conducted in Thrace to monitor the disease status of livestock with an eventual aim to demonstrate freedom from foot-and-mouth disease .

In this chapter I estimate, using a stochastic simulation model, the performance of the current Thrace serological surveys. The surveys are designed under a set of assumptions; in the analysis I relax these with regard to the animal-level diagnostic test performance and underlying epidemiology of foot-and-mouth disease. The result is an estimation of the village-level performance of the surveys. The results show that under most circumstances the surveys have sufficient sensitivity to detect disease unless present at a very low prevalence (which may not represent an important onwards transmission

threat). However the surveys' initial specificity is very poor at a village-level potentially resulting in many villages being declared as exposed to disease when they are not.

3.2 Introduction

3.2.1 Foot-and-mouth disease in Thrace

Figure 3.1: Map of North-Western Turkey with Thrace provinces named

Thrace is the European region of Turkey covering 3% of the land mass and 10% of the population (see figure 3.1). Thrace borders Bulgaria and Greece to the West and Eastern (Anatolian) Turkey to the South and East. It acts as a buffer zone between the foot-and-mouth disease endemic Anatolian region of Turkey and remainder of Europe. A European Union funded project (EUFMD Commission) assists the Turkish veterinary services with vaccination of Thrace's 403,000 cattle and 516,000 small ruminants using a trivalent foot-and-mouth disease vaccine against foot-and-mouth disease serotypes O, A and Asia 1 [DEFRA, 2005]. The short term aim of the EUFMD project is to monitor and reduce foot-and-mouth disease incursions within the area. The long term goal is permanent eradication of foot-and-mouth disease within the Thrace region and official declaration of disease freedom. Declaration of disease freedom and satisfactory disease control and monitoring systems are required as Chapter 12 of Turkey's accession to EU membership [European Council, 2008].

Movement of cattle from the foot-and-mouth disease endemic region of Anatolia into Thrace is legally banned although an annual religious festival (The Kurban Festival held on dates set according to the Islamic Calendar) results in large movements of animals into Istanbul during the festival period. Unsold animals are then allegedly moved, illegally, into Thrace maintaining a regular disease threat to the region's livestock. Foot-and-mouth disease is endemic in Anatolia (Asian Turkey) with 61 to 177 outbreaks reported each year since 2000 [DEFRA, 2005]. Two serotypes of foot-and-mouth disease are commonly found in Anatolian Turkey; Type O and Type A. Although a trivalent vaccine is used in the region, control is challenging as invasion of new subtypes mean that vaccine protection is not complete. This is a particular challenge with serotype-A which has a high antigenic diversity [Klein et al., 2006]. Some of the recent outbreaks in Anatolia have been identified as serotype-A subtype Iran 05. These may have occurred due to poor vaccination coverage or reduced cross protection with the A Iran 96 subtype on which the trivalent vaccination is based. The distribution of serotypes until 2006 in Turkey and the neighbouring Middle Eastern countries is illustrated in figure 3.2.

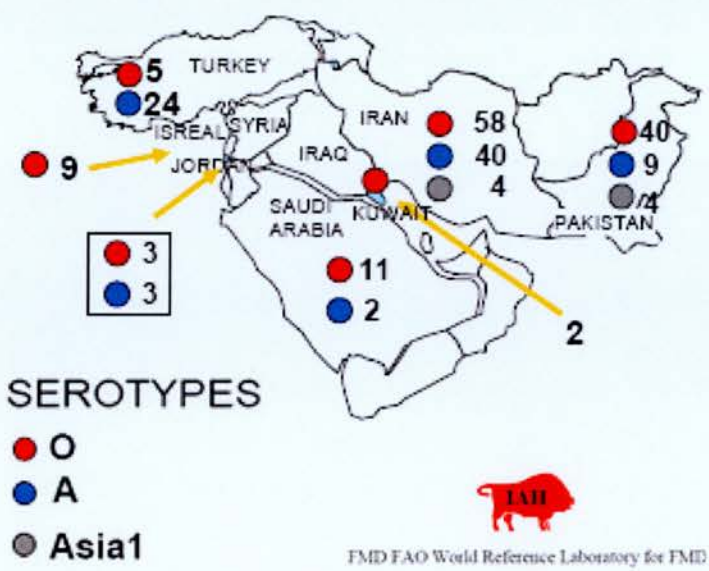


Figure 3.2: Isolates of foot-and-mouth disease from Turkey and neighbouring countries, 2004 to 2006, as identified by IAH Pirbright. (Figure from DEFRA [2007b])

Thrace has been apparently free of foot-and-mouth disease since a serotype-O outbreak in 2001 (see figure 3.3). However in early 2006 there were 16 foot-and-mouth disease outbreaks due to serotype A-22 like virus [DEFRA, 2006b]. These were thought to be the consequence of illegal movements from the Anatolia region. The outbreaks

occurred from the 11^{th} January to 26^{st} March 2006. Spring sero-surveillance samples were taken from 19^{th} February to 6^{th} March 2006. A further solitary outbreak of serotype-A Iran 22 foot-and-mouth disease occurred on the 14^{th} June 2006. Previous trivalent vaccination did not include the A-22 serotype so emergency vaccination with A-22 vaccine was carried out from 11^{th} February to 14^{th} April 2006. The estimated coverage for this campaign was 92% [EUFMD advisory group, 2006]. Outbreaks are recorded in DEFRA reports [DEFRA, 2007b,c,a, 2005, 2006a,b] (see table 3.1 for a summary from these reports).

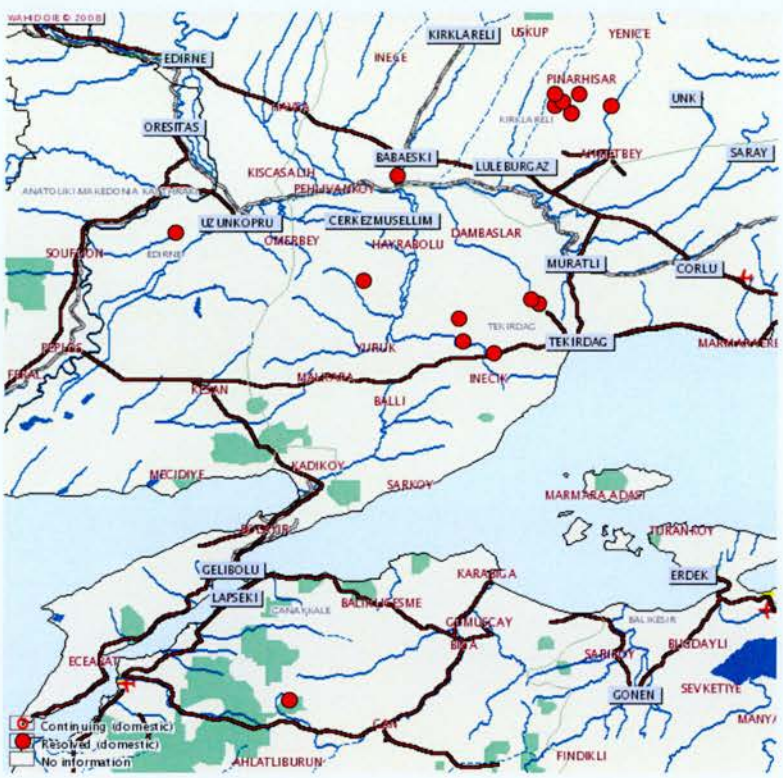


Figure 3.3: Foot-and-mouth disease outbreaks in Thrace in 2006

Province	District	Village	Date of start of outbreak	Date resolved	Cattle infected	Cattle Susceptible
Kirklareli[1,2]	Merkez	Urunlu	11/01/2006	16/03/2006	17	539
Kirklareli[1,2]	Luleburgaz	Karamusul	30/01/2006	15/03/2006	3	676
Kirklareli[2]	Babaeski	Mandira	31/01/2006		5	2306
Kirklareli[1,2]	Luleburgaz	Seyitler	01/02/2006	16/03/2006	14	125
Tekirdag[2]	Malkara	Balabancik	01/02/2006(c)		21	3354
Kirklareli[1,2]	Merkez	Kavakdere	06/02/2006	16/03/2006	17	539
Kirklareli[1]	Demirkoy	Avcilar	06/02/2006	22/03/2006	4	46
Kirklareli[2]	Dermiköy	Aveilar	06/02/2006(c)		4	46
Tekirdag[1,2]	Vakifidemir	Vakidifdemir	10/02/2006	10/03/2006	8	554
Tekirdag[1,2]	Marmaraereglisi	Y. Ciftlik	11/02/2006	14/03/2006	3	670
Kirklareli[1,2]	Merkez	Kayali	11/02/2006	16/03/2006	17	2120
Tekirdag[1]	Merkez	Cinarli	20/02/2006	21/03/2006	3	51
Kirklareli[2]	Babaeski	Nadirli	21/01/2006		4	752
Tekirdag[1]	Merkez	Kasikci	24/02/2006	23/03/2006	1	442
Tekirdag[1]	Merkez	Tasomurca	28/02/2006	24/03/2006	1	132
Edirne[1]	Uzunkopru	Kirkkavak	26/02/2006	26/03/2006	15	26
Çanakkale	(no data available)		14/06/2006		6	128

Table 3.1: Summary of reported outbreaks in Thrace region during 2006. ([1] only on OIE WHAID report, [2] only on spreadsheet from SAP Institute, Ankara, Turkey, [1],[2] on both, (c) means clinical symptoms c.f. estimated infection date). Mechanism for estimation of susceptible livestock was not recorded - may be based on vaccination census data.

International guidelines require declarations of disease freedom to be supported by a 24 month period with no foot-and-mouth disease outbreaks and a 12 month period of sero-surveillance demonstrating no circulating disease [OIE, 2009b]. To monitor effectiveness of the foot-and-mouth disease control program and eventually help demonstrate disease freedom the EUFMD Commission co-ordinate an annual (and occasionally biannual) sero-surveillance campaign aiming to detect circulating foot-and-mouth disease virus in the Thrace region.

3.2.2 Historical Surveillance Design

Currently, sero-surveillance in the Thrace region is conducted one or two times a year and has been based on OIE guidelines to substantiate freedom from disease. The surveillance uses a two-staged design with villages considered to be the epidemiological group and animals the individual unit. From a sampling frame of all the villages in Thrace, sufficient villages are sampled such that if disease were present in 2% of the villages there would be a 95% confidence that at least one infected village would be detected given the surveillance has a village-level sensitivity of 95%. Animals (cattle) are sampled from within each chosen village such that if 5% of animals within the village were infected at least one animal would test positive with a 95% confidence given a stated diagnostic test sensitivity and specificity. Currently, sero-surveillance in Thrace has uses the CEDITEST [Sorensen et al., 2005; Wen-Bin Chung et al., 2002] as the primary screening test and it has been assumed, for survey design, that this test has a sensitivity of 90% and a specificity of 100% in the Turkish cattle population.

The sero-surveys have been designed using 'Freecalc' software [Cameron and Baldock, 1998a,b] that determines a two-stage sampling design considering finite population sizes and imperfect diagnostic tests. In 2005 there were 927 villages recorded in the sampling frame. Using a herd-level sensitivity of 95% aiming for a system sensitivity or confidence of 95% a survey would need to sample from 142 villages. The current survey designs have specified 152 villages to be sampled. This permits a degree of purposive sampling of perceived higher risk areas such as Istanbul (EUFMD research group, personal communication 2006).

If two stage disease freedom surveillance, as described in Cameron and Baldock [1998b], is used as the design methodology the population of each selected village would be used

to calculate the optimal sample size for that village. However in Thrace, to avoid complicating the survey design in the field, a fixed sample size has been used within each village. This has been set to 64 animals from each village. This sample size corresponds to the 'Freecalc' optimal sample size for a population of 1000 animals with a test of 90% sensitivity and 100% specificity aiming to detect 5% prevalence with 95% confidence. With village cattle sizes below 1000 head this sample size will, theoretically, be greater than required. With larger village cattle populations the sample size may be too small although this effect will be insignificant.

Any animals testing positive with the initial CEDITEST NSP test are re-tested with a repeated CEDITEST NSP test and a BOMMELI CHEKIT NSP test [Intervet, 2008] and any animals testing negative on the second and third tests are considered negative for the purposes of substantiation of disease freedom. This strategy aims to remove false positive test results that are the consequence of diagnostic test repeatability errors. However this policy will cause a slight reduction in the overall sensitivity of the diagnosis at an animal and hence village and system level due to variance of test results in truly infected animals [Cannon, 2001].

Surveys are designed to fulfil their design criteria under a set of assumptions including disease epidemiology, livestock demography and diagnostic test performance. In this chapter I use a stochastic simulation model of the epidemiology, sampling and diagnosis of disease in the Thrace region to estimate how well the designed surveys will perform as their design assumptions are relaxed.

The epidemiological parameters describing the prevalence of disease need to be estimated/stated to evaluate serological surveillance performance. From a regulatory perspective if the rules state a design prevalence (e.g. 5% within an epidemiological group) the survey simply needs to achieve the stated confidence of detecting this prevalence. Since 2005, OIE guidelines have become less prescriptive and treat sero-surveillance as part of a portfolio of evidence to substantiate disease freedom [OIE, 2009b]. This suggests that measuring the survey systems performance against an arbitrary design prevalence/epidemiology may not be sufficient and that surveys should be assessed in their ability to detect disease present at a realistic prevalence. The challenge here is that there are limited data to estimate a *realistic* prevalence. The epidemiology of foot-and-mouth disease is dependent on demography, husbandry practice, environmental variables and

evolution of viral strains [Grubman and Baxt, 2004]. For the stochastic evaluation the surveillance performance will be assessed against both the design prevalence and a set of relatively simple but more realistic epidemiological models.

3.3 Methodology

3.3.1 Overview

Several approaches may be used to estimate the diagnostic performance of a sero-surveillance system. With simple surveys, models of foot-and-mouth disease epidemiology and test performance the overall system performance can sometimes be determined analytically [Cameron, 1997]. However for more complex survey systems and/or when it is required to model uncertainty in the knowledge of the epidemiological and diagnostic systems stochastic simulation models are normally necessary to estimate the overall diagnostic performance of the system [Tsutsui et al., 2003].

In this study a stochastic model generates simulated sets of exposed and unexposed cattle populations in villages using a set of candidate foot-and-mouth disease epidemiologies and the available cattle demography data for Thrace. Then simulated samples are drawn using the current survey designs and the analysis and interpretation of these is simulated using plausible diagnostic performance data from recent diagnostic test evaluations (see figure 3.4).

3.3.2 Demography and epidemiology

A sampling frame is available for sero-surveillance in Thrace in 2005 listing 927 villages in the region with an estimate of their cattle population (with the exception of one missing value which was excluded). The distribution of village cattle populations is shown in figure 3.5. For the simulation model, village cattle populations are randomly drawn with replacement from this empirical distribution.

A previously exposed village is modelled as having one or more exposed animals. A number of simple models can be used to describe this within-village infection distribution:

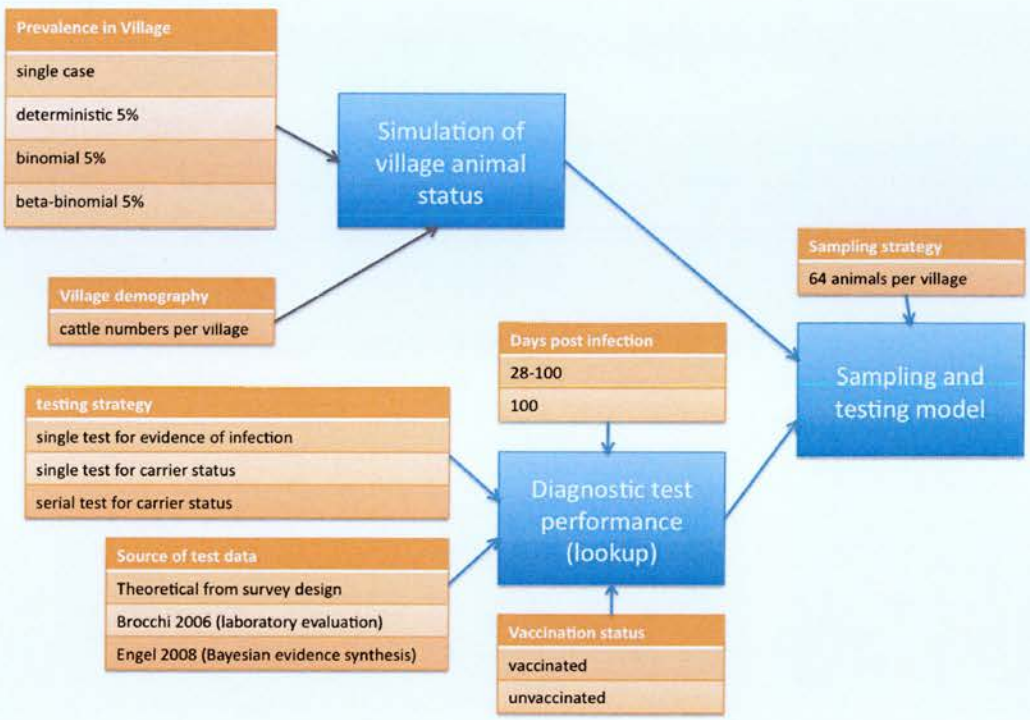


Figure 3.4: Schematic showing the components and outline structure of the stochastic evaluation model

Single individual - worst case: A single individual animal in an infected village represents the worst case from the perspective of detection of disease. Obviously this may not, necessarily, be the worst case from a disease control perspective as a single individual, especially in the generally vaccinated population of Thrace may be considered to represent a low threat of onwards transmission. However, it is appropriate to use this model as one of the scenarios for the evaluation as it sets a lower bound for the sensitivity of the village surveillance performance with respect to within-village epidemiology.

Design prevalence - deterministic: Surveillance has been historically designed in Thrace to detect infection within a village if it is present within 5% or more animals. Under this assumption 5% of each village's population is modelled as

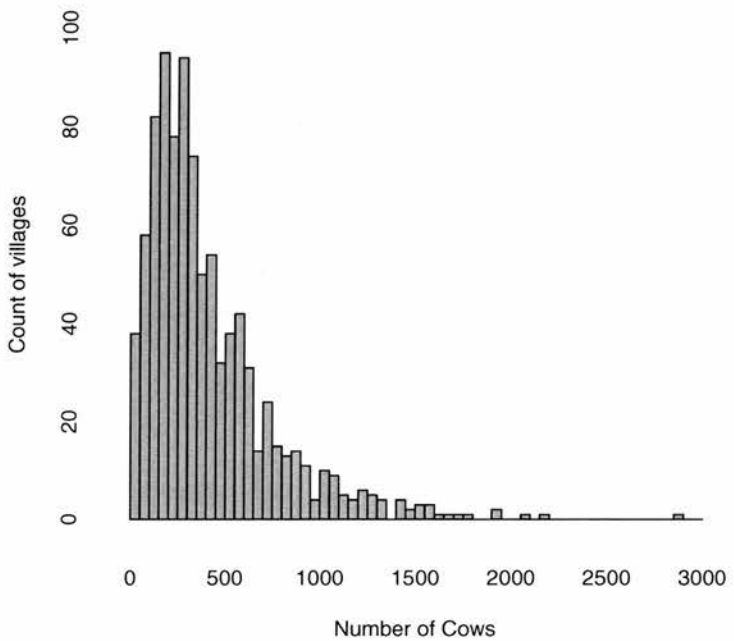


Figure 3.5: The distribution of cattle populations in Thrace villages in 2005

infected. If this value is not an integer the number of infected animals is rounded up to the next integer.

Design prevalence - binomial process: Under this assumption rather than 5% of animals in a village being infected the model assumes that the *probability* that an individual is infected is 5% and models the infection process in the village as a binomial process. This will tend to capture some of the variation in epidemic sizes between infected villages that would be seen in real outbreaks.

Design prevalence - beta-binomial process: The infection process is modelled as a binomial process within each village with a binomial probability drawn from a beta distribution for each village. The parameters of this beta distribution are set to give an overall probability of infection for an individual of 5% with an addition parameter determining the dispersion or variation in prevalence between villages [Suess et al., 2002].

The distribution of prevalence resulting from these models is shown in 3.6. These distributions represent the result of applying the above four epidemiology models to all village sizes (of greater than 64 animals — see later) and then averaging the resulting prevalence.

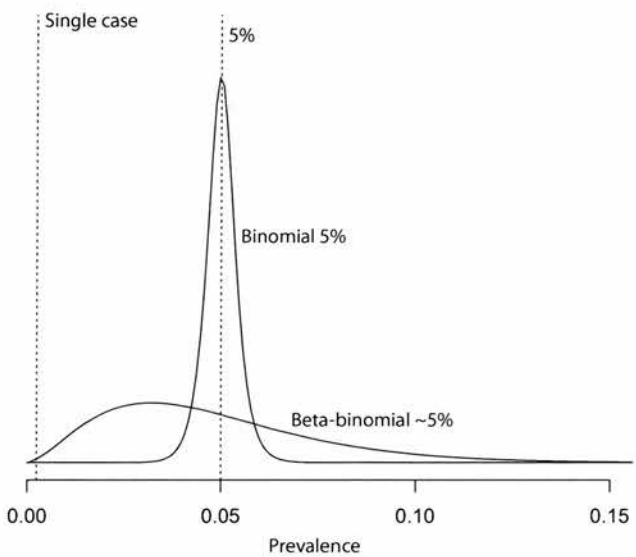


Figure 3.6: Epidemiology scenarios used in the stochastic evaluation (sampled over all herd sizes greater than 64 then smoothed)

3.3.3 Sampling strategy

Sampling in the Thrace surveillance studies has been fixed to 64 animals per village as previously discussed. In 2005, 52 of the 925 villages had a recorded cattle population of less than 64. Examination of sero-surveillance reports suggests that if a randomly selected village has a cattle population less than 64 the field veterinarians will sample animals from villages in the same district to attain the desired sample size. There are no records of the actual strategy for selecting these additional villages. For this stochastic evaluation of village-level sensitivity and specificity only villages with a cattle population of 64 or greater are considered eligible for sampling.

3.3.4 Diagnostic test performance parameters

The primary screening test is the CEDITEST ELISA test for non-structural proteins [Sorensen et al., 2005; Wen-Bin Chung et al., 2002]. For surveillance design this has been assumed to have a sensitivity of 90% and a specificity of 100%. Whilst these are convenient figures for calculation purposes the survey performance needs to be evaluated under a realistic range of diagnostic test performance parameter values. For the evaluation of survey performance I have used several models of test performance. Three derive directly from the studies by Paton et al. [2006] and Brocchi et al. [2006] where diagnostic test evaluations using animals assumed to have a known disease status are used to estimate the test's sensitivity and specificity. These evaluations will normally report a trial size (n) and the number of animals with correct diagnostic test results (r). A maximum likelihood point estimate will normally be calculated as r/n. So if 100 assumed diseased animals were tested and 98 tested positive the maximum likelihood estimate for test sensitivity would be 98%. For the stochastic evaluation an uncertainty distribution is required. For this I have used the Bayesian posterior for sensitivity and specificity assuming a Bayes prior distribution of $Beta(1, 1)$ [Lee, 1997, page 83] . This assumes that prior to the evaluation in question the belief about the test performance was uniform over 0–100%.

Design performance: A base case using 90% sensitivity and 100% specificity without modelling any uncertainty in these parameters.

Uncertain performance based on evaluations Sensitivity and specificity are considered to be uncertain and hence are sampled from distributions informed by diagnostic test evaluations in known healthy and infected animals. Uses data from Paton et al. [2006]; Brocchi et al. [2006] shown in table 3.2

Serial testing: In Thrace any animals testing positive with the initial CEDITEST screening test are then retested with a repeated CEDITEST. Paton et al. [2006] estimates the combined sensitivity and specificity of serial diagnostic test use allowing the evaluation to be extended to model the effect of this strategy. Estimates are also shown in table 3.2.

Detection of carriers: A foot-and-mouth disease carrier is defined as an animal that has an infection persistent 28 days after initial infection with virus recovered

from the oropharynx by probang sampling [Zhang and Alexandersen, 2003]. The CEDITEST has been evaluated in its ability to detect animals in the carrier state [Paton et al., 2006].

Vaccination Status	Days post infection	Infection Status	p	r	n	Posterior mean
N	28-100	Y	1.000	26	26	96.4%
N	100+	Y	0.500	1	2	50.0%
Y	28-100	Y	0.681	62	91	67.7%
Y	100+	Y	0.745	35	47	73.5%
Y	28-100	carrier†	0.864	57	66	85.3%
Y	100+	carrier†	0.892	33	37	87.2%
Y	28-100	carrier†(serial tests)	0.818	54	66	80.8%
N	-	N	0.998	656	675	97.0%
Y	-	N	0.995	423	425	99.3%
Y	N	N (serial tests)	0.995	423	425	99.3%

Table 3.2: Parameter estimates of CEDITEST performance for single testing strategy (from [Brocchi et al., 2006; Paton et al., 2006]) (p is reported sensitivity or specificity, n is reported sample size and r is inferred number of test positive (or negative) animals for infected (or uninfected) subjects). († animals from which virus is isolated more than 28 days after infection)

A further two models use the diagnostic test performance estimates of Engel et al. [2008]; this evaluation relaxes the assumption that the disease status of the sampled animals is known and use the joint test results of six different NSP diagnostic tests to jointly estimate the sensitivity and specificity of the tests and the prevalence of disease in the test subject groups. This approach derives from the Hui-Walters methodology of assessing diagnostic tests without having a gold-standard and, potentially, provides less biased estimates of test performance with more accurate estimates of the uncertainty of the performance estimates [Hui and Walter, 1980]. The results of Engel et al. [2008]'s study are distributions for the sensitivity and specificity of the diagnostic tests. In a fully Bayesian analysis it would be possible to repeat Engel et al. [2008]'s model to produce samples from the joint posterior distributions of sensitivity and specificity for the CEDITEST. This approach does have high computational requirements so as an alternative I have used beta distributions to approximate the results from Engel et al. [2008]'s study. Their paper records means and 0.025 and 0.975 quantiles for each diagnostic test performance measure. I estimated Beta distribution parameters to

approximate these results using a Nelder-Mead optimiser ('optim' function [R Development Core Team, 2007]) to minimise the sum of squared differences between three statistics of the reported and candidate distributions:

$$\sum \left[(\bar{t} - \bar{c})^2 + (q_{0.025t} - q_{0.025c})^2 + (q_{0.975t} - q_{0.975c})^2 \right]$$

where q_n is the n^{th} quantile, t is the statistic from the study and c is the statistics from the approximating distribution. Estimates for the beta parameters are shown in table 3.3.

Vaccination Status	Age	Infection Status	a	b	mean
N	-	N	738.7	22.2	97
N	-	Y	52.3	1.7	97
Y	-	N	593.9	6.6	99
Y	-	Y	141.3	25.3	85

Table 3.3: Parameter estimates of CEDITEST performance for single testing strategy (from Engel et al. [2008]) parameters a and b from fitted Beta(a,b)

In Brocchi et al. [2006] diagnostic sensitivity was assessed at a number of periods after infection. For my evaluation the periods of 28-100 and greater than 100 days were used; the former as it reflects likely post epidemic surveillance timing and the latter to cover the possibility that surveillance (which is annually to twice annually) may take place more than 100 days after exposure/infection events.

Since 2000 EUFMD [1999] cattle in Thrace have been vaccinated, biannually against foot-and-mouth disease. Both the epidemiology of disease within an infected group of animals and the performance of a diagnostic test will vary in response to vaccination. Hence the survey evaluation needs to model the influence of vaccination on these parameters and therefore an estimate of vaccination coverage is required. There are limited data available to estimate coverage. A survey dated March 2006 [Bulut, 2006] reports an overall vaccination coverage for Thrace of approximately 90%. There is a wide variation in estimated coverage across the districts with some districts apparently having received no vaccinations and others with coverage exceeding 100% (in one case 20% more cattle were vaccinated than were thought to be in the district). Hence these data can only be used to approximate coverage levels. For this evaluation of survey performance village-level sensitivity and specificity will be assessed using diagnostic test evaluation results from studies in unvaccinated and vaccinated cattle.

3.3.5 Detailed model Structure

The model estimates the herd-level performance of the survey system. This is estimated by simulating the testing and interpretation of results for a large number of exposed (to estimate village-level sensitivity) and unexposed (to estimate village-level specificity) villages. The village sizes are drawn from the 2005 census data so that the performance estimates are representative of the performance of the survey in a randomly selected village.

For exposed villages the number of infected animals in each simulated village will be drawn from one of the four candidate epidemiologies:

Single animal $n_{exposed} = 1$ animal infected in each village

Design prevalence - deterministic 5% of animals infected in each village $n_{exposed} = ceiling(h \times 0.05)$

Design prevalence - stochastic $n_{exposed} \sim Binomial(h, 0.05) + 1$ animals infected where h is the village size

Beta binomial distribution $n_{exposed} \sim Binomial(h, Beta(2.9, 58.1)) + 1$ animals infected where h is the village size

Where h is the village cattle population and $ceiling(x)$ is a function returning the lowest integer greater than x

In the beta-binomial model the beta parameters were selected to give a mean prevalence within the 2005 herd size distribution of 5% (i.e. identical to the design prevalence) but to introduce over-dispersion compared to a binomial distribution with a 95% probability interval of approximately 0.01 to 0.1.

For each simulation for each exposed village 64 samples are drawn without replacement so the number of exposed animals is given by:

$$s_{exposed} \sim Hyper(n_{exposed}, h - n_{exposed}, 64)$$

Where $Hyper(n, m, k)$ is the hypergeometric distribution function of number of exposed animals drawn for a sample (without replacement) of size k from a group of animals where n are exposed and k are not.

The number of unexposed animals in each sample is simply:

$$s_{unexposed} = 64 - s_{exposed}$$

The model then simulates the testing of each sample with a potentially imperfect diagnostic test so that the number of test positive samples is given by:

$$t_{positive} \sim Binomial(s_{exposed}, Se) + Binomial(s_{unexposed}, (1 - Sp))$$

and test negative samples by:

$$t_{negative} = 64 - t_{positive}$$

Where Se and Sp are the test sensitivity and specificity of the CEDITEST ELISA test respectively. Values for these two parameters are draw from the relevant distributions for each village sampling/testing simulation as listed in section 3.3.4

To simulate sampling and testing of an unexposed herd the number of test positives is simply:

$$t_{positive} \sim Binomial(64, (1 - Sp))$$

Thus the model gives the number of test positive and test negative animals for each simulation of either an exposed or unexposed herd under the different epidemiological and diagnostic test scenarios. In traditional sero-surveillance to demonstrate disease freedom the result of testing a village would then be considered with respect to a village cut-point such that observation of more test-positive animals than this cut-point would result in the village being classified as exposed. In this analysis I use a cut-point of zero as this is will maximise village-level specificity. Hence any village with one or more test positive animals is regarded as exposed. Over the full set of simulations for each

combination of epidemiology and diagnostic test scenarios the village-level sensitivity and specificity can be estimated using:

$$VSe = \frac{\sum(V^+)}{nSim_{exposed}}$$

$$VSp = \frac{\sum(1 - V+)}{nSim_{unexposed}}$$

Where V^+ is an indicator variable which is 1 if a village is classified as exposed and 0 if it is not and $nSim$ is the total number of simulations for the exposed or unexposed simulation sets.

The simulation model estimates the sensitivity and specificity of the surveillance system at a village-level. There are several considerations that need to be applied to the interpretation of these results:

3.3.6 Single test analysis as an upper bound of group and system level performance

In a practical testing situation any animals testing positive may be subsequently retested with a confirmatory test. Serial testing strategies then require this second test to be positive for the animal and hence village to be classified as infected. The upper bound of the sensitivity of such surveys at village-level is defined by the application of the first test i.e. subsequent tests to rule out false positive animals will only *reduce* the overall sensitivity of the system at village-level. Thus for this analysis the estimated sensitivity results from single test applications can be considered an upper bound for the village-level sensitivity.

3.3.7 The effect of imperfect specificity on group level sensitivity

At a village or regional level the sensitivity of a survey to substantiate disease freedom will be increased if the animal-level test is not perfectly specific. A test with less than 100% specificity may result in some animals being classified as infected when they are not. This will then result in the classification of some infected villages of animals as

infected even though no infected animals were sampled and classified as infected by the diagnostic test. Although this anomaly apparently increases the village-level sensitivity of the survey it is problematic. Villages found to be infected on the first survey will normally be subject to follow up investigation involving re-testing of the submitted samples. A more specific test would correctly identify these samples as from uninfected animals and thus the village would be finally reclassified as uninfected.

To remove this effect to allow assessment of the sensitivity of the surveillance systems arising solely as a consequence of the diagnostic test's sensitivity all my evaluations include an analysis where the the diagnostic test specificity is set to 100%.

3.4 Results

Results of the stochastic simulation model are shown in tables 3.4 – 3.13

Using test performance assumed for current survey design

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.22	0.22	1.00
2	Design 5% prevalence	0.96	0.96	1.00
3	Binomial 5% probability	0.95	0.95	1.00
4	Beta-binomial ~5% probability	0.89	0.89	1.00

Table 3.4: Estimation of village-level performance of surveillance **Design test performance** Village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

The estimated survey village-level performance with the design test performance of 90% sensitivity and 100% specificity is shown in table 3.4. The surveys are designed to have a target of 95% village-level performance; the estimated performance using a 5% binomial model for infection of 95% sensitivity is as expected. If a fixed village by village prevalence of 5% is used (rounding up to the next highest integer) the village performance is slightly higher at 96%. If prevalence is allowed to vary between villages, though still averaging 5% the average sensitivity performance is 89%, markedly worse than the target performance. Effectively, the improved performance in villages having

higher prevalence is not sufficient to compensate for the degraded performance in villages having a lower prevalence. This is a consequence of the village-level performance being a non-linear function of the prevalence as shown in figure 3.7. In the worst case scenario - where only a single individual in each village is infected, the village-level performance is drastically worse than the target performance with a sensitivity of 20%.

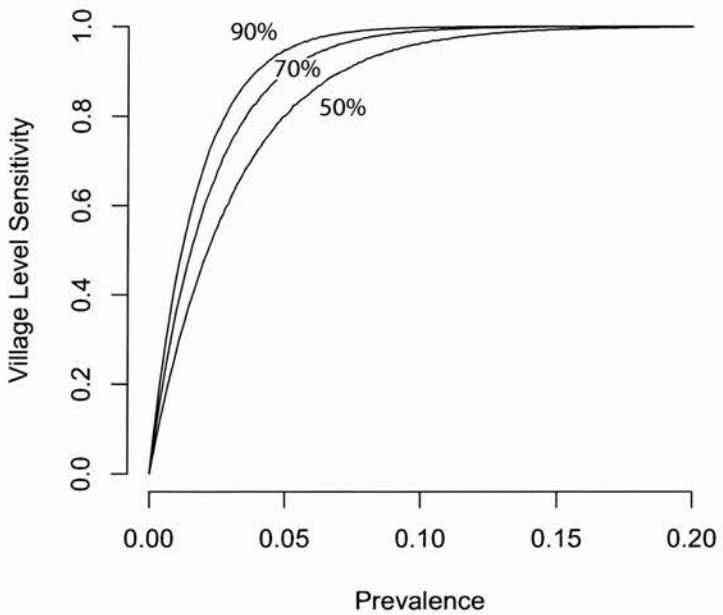


Figure 3.7: The non-linear effect of animal-level prevalence of disease on village-level sensitivity. The curves show the estimated village-level sensitivity of disease detection for a range of prevalence using tests of 50%, 70% and 90% animal-level diagnostic sensitivity. The curves were generated using simulated sampling of 64 animals from a village of 500 animals. A binomial model of infection was used within the village.

The village-level specificity is 100% as expected as the diagnostic test is modelled to have perfect specificity.

Vaccinated cattle using estimates from Brocchi 2006 study

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.45	0.16	0.66
2	Design 5% prevalence	0.94	0.91	0.66
3	Binomial 5% probability	0.93	0.89	0.66
4	Beta-binomial ~5% probability	0.89	0.83	0.65

Table 3.5: Estimation of village-level performance of surveillance in **vaccinated cattle at 28–100 days post infection**. Village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.47	0.17	0.65
2	Design 5% prevalence	0.95	0.93	0.66
3	Binomial 5% probability	0.94	0.91	0.66
4	Beta-binomial ~5% probability	0.90	0.85	0.65

Table 3.6: Estimation of village-level performance of surveillance in **vaccinated cattle at 100+ days post infection**. Village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

Tables 3.5 and 3.6 show the estimated performance using diagnostic test performance data from Brocchi et al. [2006] in animals 28-100 and 100+ days post infection. These post infection timings represent the likely post infection times of disease surveillance testing so the real world village-level performance is likely to lie somewhere between the estimates in these two tables.

In practice the results for 28-100 and 100+ days post infection are similar. This is a consequence of similar estimates of diagnostic test performance. Using the village-level sensitivity estimates assuming perfect diagnostic test specificity the village sensitivity is less than 95% for all epidemiology scenarios. With the fixed and binomial models the village-level sensitivity is 90% but this drops to 85% if the prevalence is allowed to vary between villages (but has a mean value of 5%). In the extreme case of a single exposed animal in a village the sensitivity is 16-17%.

The village-level specificity is about 65% in both diagnostic test scenarios. This means that on average about one in three disease free villages tested will initially have at least one test positive animal and will require further follow up testing.

Unvaccinated cattle using estimates from Brocchi 2006 study

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.88	0.24	0.16
2	Design 5% prevalence	1.00	0.97	0.16
3	Binomial 5% probability	0.99	0.96	0.16
4	Beta-binomial ~5% probability	0.99	0.90	0.16

Table 3.7: Estimation of village-level performance of surveillance **Un-vaccinated cattle at 28–100 days post infection** village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.86	0.12	0.16
2	Design 5% prevalence	0.96	0.76	0.16
3	Binomial 5% probability	0.96	0.75	0.17
4	Beta-binomial ~5% probability	0.95	0.69	0.16

Table 3.8: Estimation of village-level performance of surveillance **Un-vaccinated cattle at 100+ days post infection**village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

Tables 3.7 and 3.8 show the estimated performance at a village-level in unvaccinated animals using estimates of diagnostic test performance from Brocchi et al. [2006]. These diagnostic test performance parameters give specificity corrected village-level sensitivity of 96-98% for design and binomial scenarios in 28-100 days post infection animals. In the beta-binomial model the sensitivity in these animals is lower at 90%. In the worst case scenario of single animal infections the village-level performance is only 24%. Village-level specificity reflects the lower diagnostic test specificity and is approximately 16%

with 5 out of 6 villages being reported as exposed after the initial application of the testing strategy.

The village-level sensitivity results in animals 100+ days post infection are all low. This is a consequence of the low estimate of diagnostic test sensitivity in this group (point estimate 50% but with large uncertainty as a very small group was used to estimate the parameters). In these animals the highest village-level sensitivity is 76% in the design scenario.

Vaccinated cattle using estimates from Brocchi 2006 study — detecting carriers at 28–100 days

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.49	0.20	0.67
2	Design 5% prevalence	0.97	0.96	0.66
3	Binomial 5% probability	0.96	0.94	0.66
4	Beta-binomial ~5% probability	0.92	0.88	0.65

Table 3.9: Estimation of village-level performance of surveillance **Vaccinated cattle at 28–100 days post infection to identify carrier state** Village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

As an alternative to assessing the surveillance scheme in its ability to detect exposed animals (using experimentally infected animals as a proxy for this in the diagnostic test evaluation data selection) we can consider detection of carrier animals. Brocchi et al. [2006] classified an animal as a carrier if there was evidence of persistent infection more than 28 days post initial infection. Virus isolation or positive RT-PCR were considered signs of infection. The data for carrier detection performance were modelled in vaccinated animals at 28-100 post infection and is shown in table 3.9. The specificity corrected village-level sensitivity is close to the target 95% level for the design and binomial scenarios and below (88%) for the beta-binomial scenario. In the single case scenario the expected performance is 20%.

Vaccinated cattle using estimates from Brocchi 2006 study — detecting carriers at 100+ days

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.48	0.21	0.65
2	Design 5% prevalence	0.97	0.96	0.66
3	Binomial 5% probability	0.96	0.94	0.65
4	Beta-binomial ~5% probability	0.92	0.88	0.66

Table 3.10: Estimation of village-level performance of surveillance **Vaccinated cattle at 100+ days post infection to identify carrier state**Village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

The results for carrier detection at 100 days+ are shown in table 3.10. The CEDITEST diagnostic test performance is slightly higher in these animals at 87%. The village-level results for specificity corrected sensitivity are similar to the 28-100 day post infection carriers.

Vaccinated cattle using estimates from Brocchi 2006 study — detecting carriers at 28–100 days using serial testing

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.47	0.20	0.66
2	Design 5% prevalence	0.96	0.95	0.66
3	Binomial 5% probability	0.95	0.93	0.66
4	Beta-binomial ~5% probability	0.91	0.87	0.66

Table 3.11: Estimation of village-level performance of surveillance **Vaccinated cattle at 100+ days post infection to identify carrier state with serial testing**. Village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

Serial testing is where any positive results on an initial CEDITEST are followed up with a second test which must also be positive to declare an animal infected. This will theoretically increase specificity and reduce sensitivity [Gardner et al., 2000]. Unless the tests are completely conditionally dependent, disease-free animals testing positive

to the first test will have a some probability of testing negative to the second test — increasing specificity. Diseased animals testing positive to the first test will likewise, as long as the tests are not conditionally dependent, have some possibility of testing negative to the second test — reducing sensitivity. In the Brocchi et al. [2006] evaluations the CEDITEST diagnostic test specificity estimate remained the same (within the evaluation precision) at 99.5% and sensitivity was reduced to 81%. The village-level results are shown in table 3.11. Specificity corrected village-level sensitivity was hence reduced by approximately 1% point for each scenario.

Unvaccinated cattle using estimates from Engel 2008 study

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.87	0.24	0.16
2	Design 5% prevalence	1.00	0.97	0.16
3	Binomial 5% probability	0.99	0.96	0.16
4	Beta-binomial ~5% probability	0.99	0.91	0.16

Table 3.12: Estimation of village-level performance of surveillance **Unvaccinated cattle using Engel Bayesian estimates of diagnostic test performance**Village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

Using the estimates of diagnostic test performance from Engel et al. [2008] the corrected village-level sensitivity exceeded the target level for the design and binomial scenarios. It dropped to 91% for the beta binomial scenario and was 21% for the single case scenario. Village-level specificity was approximately 16%.

Vaccinated cattle using estimates from Engel 2008 study

Using the estimates of diagnostic test performance from [Engel et al., 2008] the corrected village-level sensitivity was close to the target level for the design and binomial scenarios. It dropped to 88% for the beta binomial scenario and was 21% for the single case scenario. Village-level specificity was approximately 51%.

	Epidemiology	HSe	HSe2	HSp
1	Worst case - single infected	0.58	0.21	0.50
2	Design 5% prevalence	0.98	0.95	0.51
3	Binomial 5% probability	0.97	0.94	0.50
4	Beta-binomial ~5% probability	0.94	0.88	0.51

Table 3.13: Estimation of village-level performance of surveillance **Vaccinated cattle using Engel Bayesian estimates of diagnostic test performance.** Village-level sensitivity (HSe), village-level sensitivity assuming perfect specificity (HSe2) and village-level specificity (HSp)

3.4.1 Summary

The results are also shown sorted by epidemiology model for corrected village-level sensitivity in table 3.14 and in figure 3.8.

Epidemiology	Vaccination/Testing	Days post infection	Study	HSe2
Single animal	-		design	21.92
	un-vacc	28-100	brocchi	23.58
	un-vacc	100+	brocchi	12.24
	un-vacc		engel	23.69
	vacc	28-100	brocchi	16.61
	vacc	100+	brocchi	17.98
	vacc		engel	20.75
	vacc carrier	28-100	brocchi	20.76
	vacc carrier	100+	brocchi	21.30
	vacc carrier serial test	28-100	brocchi	19.78
5% deterministic	-		design	96.34
	un-vacc	28-100	brocchi	97.18
	un-vacc	100+	brocchi	76.18
	vacc		engel	95.43
	vacc	28-100	brocchi	90.84
	vacc	100+	brocchi	92.64
	un-vacc		engel	97.28
	vacc carrier	28-100	brocchi	95.50
	vacc carrier	100+	brocchi	95.81
	vacc carrier serial test	28-100	brocchi	94.63
5% Binomial	-		design	94.96
	un-vacc	28-100	brocchi	95.92
	un-vacc	100+	brocchi	74.63
	un-vacc		engel	96.06
	vacc	28-100	brocchi	89.03
	vacc	100+	brocchi	90.90
	vacc		engel	93.95
	vacc carrier	28-100	brocchi	93.98
	vacc carrier	100+	brocchi	94.39
	vacc carrier serial test	28100	brocchi	93.04
5% Beta-binomial	-		design	89.15
	un-vacc	28-100	brocchi	90.37
	un-vacc	100+	brocchi	69.83
	un-vacc		engel	90.47
	vacc	28-100	brocchi	82.81
	vacc	100+	brocchi	84.72
	vacc		engel	87.96
	vacc carrier	28-100	brocchi	88.00
	vacc carrier	100+	brocchi	88.46
	vacc carrier serial test	28-100	brocchi	86.88

Table 3.14: Summary estimates of village-level sensitivity (assuming diagnostic test has perfect specificity) grouped by epidemiology model and ordered by sensitivity. Study refers to source data for diagnostic test performance(design – 90% Se and 100% Sp, Brocchi – [Brocchi et al., 2006], Engel – [Engel et al., 2008]. Vaccination/testing refers to vaccination status and additionally if ability to detect carriers was the assessment criterion (with single or serial testing))

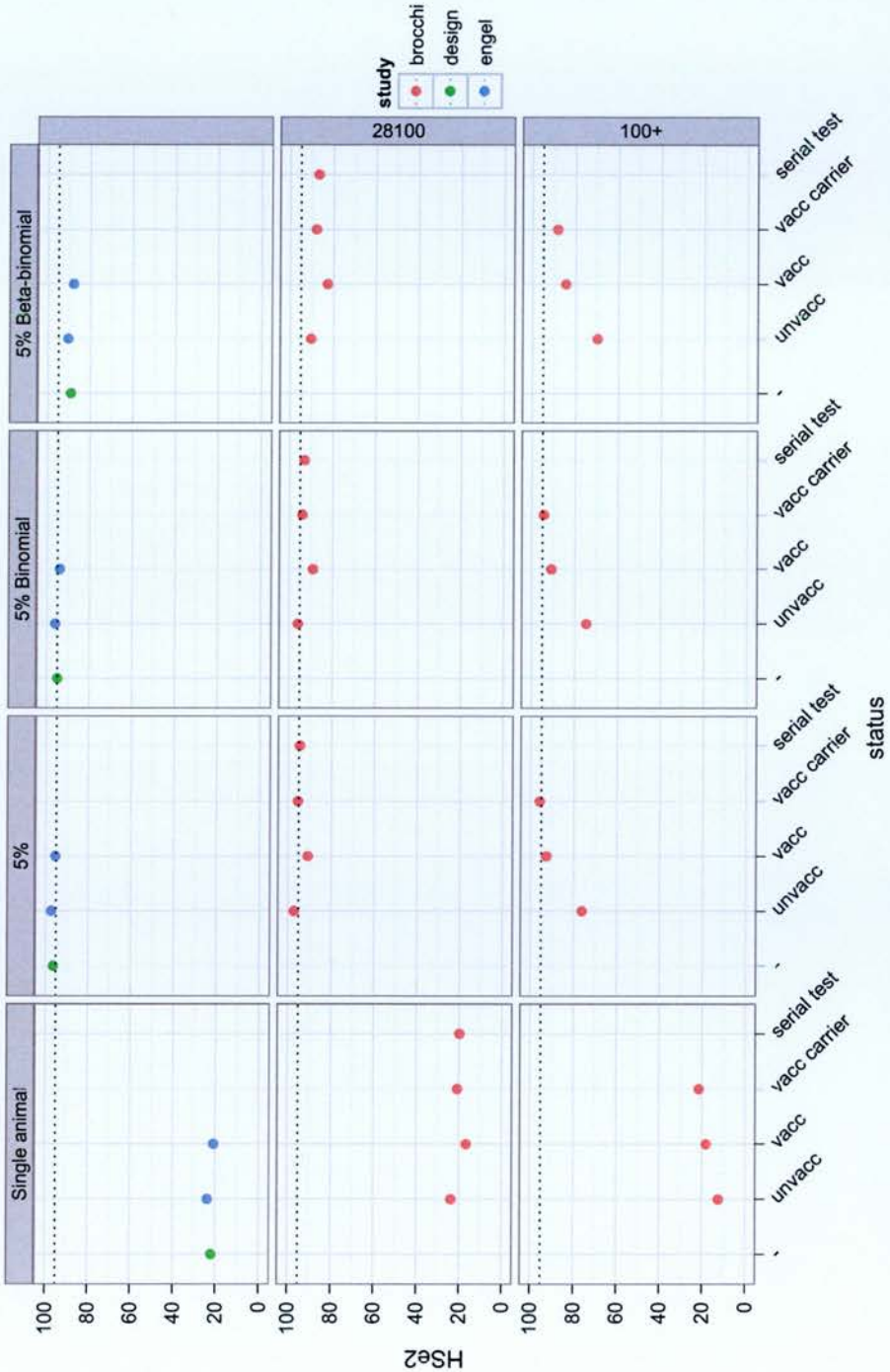


Figure 3.8: Village-level sensitivity assuming perfect diagnostic test specificity. 95% village level sensitivity is shown with a dotted horizontal line. (Status is vaccination/carrier type detected by test. Study is source of CEDITEST performance data. Plots are tiled vertically by days post infection and horizontally by epidemiology scenario)

3.5 Discussion

The performance of a disease surveillance system is contingent on the true epidemiology of the disease, the sampling strategy, the diagnostic test performance and the decision rule for the interpretation of the consequent results. In this analysis I have used worst case and design based estimates of epidemiology. The sampling strategy is that previously used in the Thrace sero-surveillance exercises. The diagnostic test performance estimates are those available from a large, quality controlled test evaluation in addition to using the values of test performance used to design the sero-surveillance in Thrace.

For all estimates of test performance other than the design test performance (with 100% test specificity) the imperfect specificity of the diagnostic test gave an apparent increase in the village-level sensitivity. False positive animal results in villages where no diseased animals were sampled and diagnosed diseased resulted in correct diagnosis of village status but for an incorrect reason. This form of increased sensitivity is not valuable as the confirmatory testing of test positive animals would tend to demonstrate their true, non-diseased, status and the initial higher estimate of village-level sensitivity would be very optimistic. Hence for the remainder of this discussion I will only consider the specificity corrected sensitivity (i.e. assuming that any test positive animals are followed up with tests of perfect sensitivity and specificity). This value gives a realistic upper bound for the overall system sensitivity with followup.

3.5.1 Sampling from smaller villages

In this evaluation I have restricted the simulation to villages holding 64 or more cows. In the Thrace field surveys if a randomly selected village has less than 64 cows sampled a neighbouring village will be used to make up the sample numbers to 64. Disease freedom surveillance in smaller animal groups has been described in detail by Greiner and Dekker [2005]. The main design issue occurs when the diagnostic test sensitivity is lower than the target sensitivity require at the group (in this case village) level. With a sufficiently large group enough samples can be taken to reach the target sensitivity for the group. However, with villages smaller than 64 animals there will be a compensatory effect as *all* animals will be sampled in the primary sampling group; even in a worse case where there is only one diseased animal, it will be tested. This advantage will not be true of the subsequent villages selected to top-up the samples to 64. It is likely

that the presence of disease in these, 'top-up' villages used to supplement the small villages will be correlated with the presence in the primary sampling village as foot-and-mouth disease will spread readily by local movements, airborne spread and fomite spread [Alexandersen et al., 2003]. Villages with less than 64 cattle represent 5.6% of the 2005 sampling frame (52 of 924 villages) so in the absence of adequate spatial models of disease in Thrace and limited knowledge of the sampling behaviour I consider the exclusion of these smaller villages provides a reasonable estimate of overall expected village-level performance.

3.5.2 Performance in worst case scenarios

The current surveillance design does not achieve a village-level sensitivity of 95% when a single animal is diseased under any of the plausible diagnostic test estimates. Although this is not surprising given the design prevalence for the survey it is potentially of concern as the Thrace cattle population is largely vaccinated. Vaccination will reduce the transmission probability of foot-and-mouth disease [Orsel et al., 2005, 2007a] and hence result, on average, in smaller within group prevalences. Furthermore, clinical detection may be less likely due to reduced clinical signs in vaccinated animals. Hence to use a sampling based approach to make strong statements about disease status at a village-level will require convincing evidence about likely village-level prevalences (if disease were present) in the face of the current vaccination coverage.

3.5.3 Performance in 5% prevalence scenarios

In the scenario with within-village prevalence modelled as a deterministic 5% the surveys will generally detect disease in unvaccinated animals with a sensitivity close to or greater than 95% with the exception of detection in unvaccinated animals 100+ days post infection. The CEDITEST evaluation data for this particular group was based on a sample of only 3 animals. The stochastic evaluation results reflect this uncertainty but it is possible that the performance estimates for this group would improve with more precise CEDITEST evaluation data. Indeed the estimates using the data from Engel et al. [2008] the performance for unvaccinated cattle using the 5% deterministic scenario is approximately 97%.

In vaccinated cattle the village-level sensitivity was consistently below the OIE target figure of 95% using the estimates of test performance from the Brocchi et al. [2006] studies. Using data test performance estimates from the Engel et al. [2008] evaluation the village-level performance generally reached 95% apart from in the beta-binomial prevalence model scenario. These results in vaccinated cattle are of particular concern as prevalence of foot-and-mouth disease may be lower in vaccinated villages due to reduced transmission probability.

The beta-binomial epidemiology scenario gave an expected village-level sensitivity of less than 95%. This is a concern as this scenario may be more likely to model the real life situation with inter-village variation of prevalence. As the village-level sensitivity is a non-linear function of prevalence an average prevalence of 5% within villages using a design for deterministic prevalence of 5% does not give a design performance.

3.5.4 Concluding remarks

In all but the worse case scenarios and the use in unvaccinated animals more than 100 days post infection the village-level sensitivity was greater than 80%. Over the 152 or more village-level samples taken each year this will still give a regional level sensitivity performance of at least 92.3% (estimated using FreeCalc [Cameron and Baldock, 1998a,b]) however although this allows region level decisions to be made with high confidence the survey system may be less sensitive at a village-level than intended. The survey has practical issues related to the village-level specificity. This will result in a high rate of false positive villages if a single positive animal is used as a cut-point. Although confirmatory testing will ultimately allow a large proportion of these villages to be correctly classified this procedure is costly, time consuming and will not remove all false positive villages. An alternative strategy is to increase focus on disease control with efforts to achieve uniformly high coverage vaccination and to increase biosecurity thus reducing the frequency and magnitude of foot-and-mouth disease outbreaks. Serological surveillance could focus on estimation of outbreak magnitude by application of NSP tests and confirmatory testing of whole animal groups only after clinical detection. This would consume fewer resources allowing precise estimation of the benefits of vaccination in controlling outbreaks after disease incursions.

Chapter 4

Thrace sero-surveillance: Exploratory data analysis

4.1 Summary

Survey data are available for three surveys (one carried out in 2005 and two in 2006). In this chapter I describe the structure of these surveys and summarise the results. The number of animals testing positive over the whole survey increased from 2005 to 2006. This is coincident with a number of village outbreaks of foot-and-mouth disease which occurred in Spring and Summer of 2006 starting at about the time of the second survey. The diagnostic test used in these surveys for initial testing (the CEDITEST NSP) produces a continuous numerical result for each test which is dichotomised into a *positive* or *negative* classification by comparison to a cut-off value. In this exploratory analysis there is evidence of a shift in the continuous results between the two surveys in 2006. This may result in misclassification of animals disease exposure status. I suggest reasons for this and propose a method by which the test can be retrospectively re-calibrated to attempt to correct for the shift.

4.2 Introduction

The background to the foot-and-mouth disease situation in Thrace is described in chapter 3. In this chapter the data from three serological surveys is described and

reported through exploratory data analysis prior to the modelling of these data-sets in chapter 5. The data-sets describe the laboratory results from three serological surveys carried out in 2005 and 2006. The timing of the three surveys and the approximate times of the Thrace foot-and-mouth disease outbreaks are shown in figure 4.1.

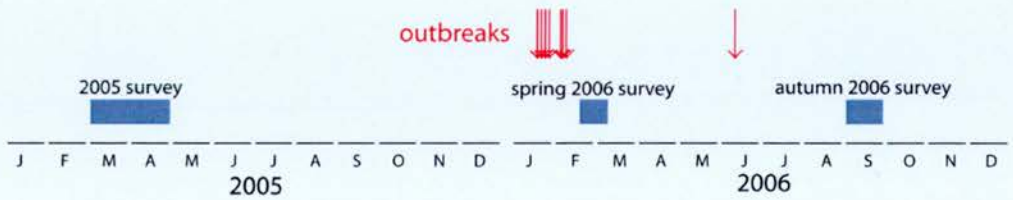


Figure 4.1: A summary of the timing of sero-surveillance in Thrace (2005 and 2006) showing approximate timing of Thrace foot-and-mouth disease outbreaks

4.3 Methodology

Survey data and sample frame data were obtained from Excel spreadsheets supplied by the ŞAP foot-and-mouth disease institute in Ankara Turkey. The spread sheet data was cleaned by:

- Removal of blank lines and additional fields
- Removal of any records with null or zero results - assumed to be non-analysed
- Re-coding of test results from Turkish to English (pos/neg etc)
- Relabelling of field headers from Turkish to English

The spreadsheet data was then manually imported into a Microsoft Access database (Microsoft, Redmond CA). SQL queries were then used to aggregate and summarise the data by provinces, villages and surveys as required for the exploratory analysis. Query results were then exported into the R statistical system [R Development Core Team, 2007] for exploratory data analysis and plotting.

4.4 Results

4.4.1 Data Sources

Animal health and surveillance records for the Thrace region are structured within the following administrative/stock holding hierarchy:

1. Province (5 total)
2. District (30 total)
3. Village (926 total)
4. Farm (approximately 1-10 per village)
5. Stock (Cattle/Sheep/Goats)

4.4.2 Key to tables and figures

2005 Spring 2005 Survey

2006A Spring 2006 Survey

2006B Autumn 2006 Survey

4.4.3 Sample frame - Thrace demography

In the 2005 sample frame data there are 5 provinces (Edïrne, Kirklareli, Tekïrdağ, Çannakale and İstanbul). These provinces are composed of administrative districts within which are villages which will have a number of farmers owning livestock. The numbers of districts, villages and cattle within each province are shown in table 4.1 together with the mean and standard deviation of the cattle populations in each province's villages. The distribution of recorded cattle populations within these villages is shown, by province, in figure 4.2. The agricultural census records do not contain data describing the distribution of livestock holdings at an owner/farm level.

Table 4.1: Number of districts, number of villages, number of cattle, mean and standard deviation of village cattle population in each province as recorded in 2005 census

	Edïrne	Kirklarelï	Tekïrdağ	Çannakale	İstanbul
Population					
Districts	9	8	9	2	4
Villages	278	202	288	42	116
Cattle	128,953	80,035	105,640	10,807	56,641
Cattle per village					
Mean	463.86	396.21	365.54	257.31	488.28
SD	330.80	286.81	357.12	281.49	404.61

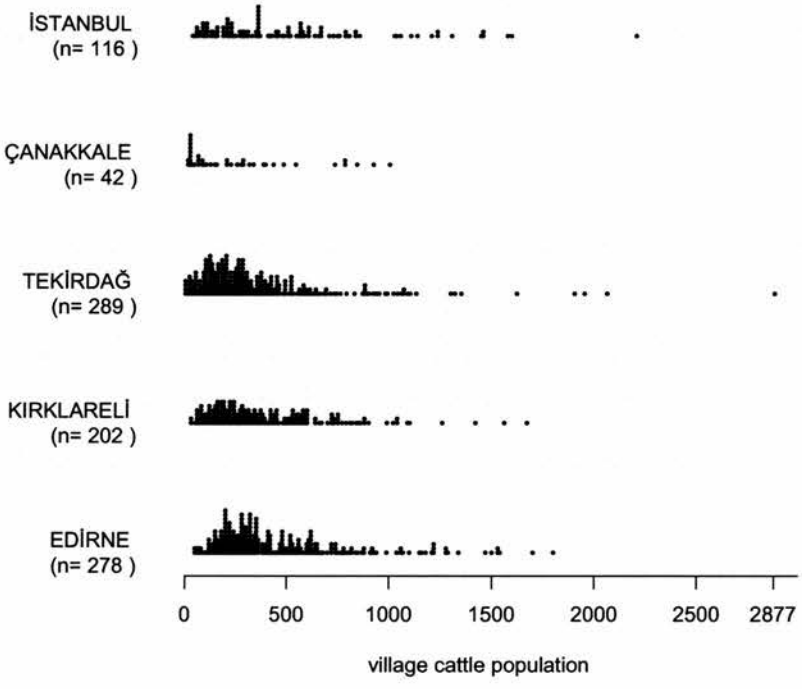


Figure 4.2: Village cattle population sizes (2005 census data) by province. Each village's cattle population is represented by a black dot. Villages with similar cattle populations in each province have their dots stacked in this plot.

4.4.4 Dichotomous results from sero-surveillance

Sero-surveillance results are available from three separate surveys conducted in 2005, Spring 2006 and Autumn 2006 as shown in figure 4.1. The number of villages visited in each survey are shown in table 4.2. In all surveys the survey design required 64 animals to be sampled in each village. However some villages did not have 64 cattle available for testing so the district veterinary surgeons occasionally needed to sample animals from neighbouring villages. The resulting distribution of cattle numbers sampled by village is shown in 4.3. Occasionally villages were over-sampled with 6 villages in Spring 2006 being double sampled. In all surveys some villages had more than 64 samples taken. There are no field records or documentation available to give background to this over-sampling behaviour.

Table 4.2: Number of villages sampled in each survey

Survey	Villages sampled
2005	162
2006 Spring	239
2006 Autumn	196

Table 4.3: Samples taken by village in each survey

Samples	4	6	8	11	13	14	25	26	38	39	50	53	63	**64**	65
2005	1	1	2	2	1	2	1	1	1	1	1	1	1	**145**	1

Samples	8	10	12	17	34	41	47	60	**64**	87	128
2006 Spring	1	1	1	1	1	1	1	1	**224**	1	6

Samples	23	26	28	38	**64**	77
2006 Autumn	1	1	1	1	**191**	1

Samples were initially analysed by the ŞAP institute in Ankara with a single CEDITEST NSP test [Sorensen et al., 1998] using a 50% cut-off in in the percentage inhibition result to classify the cattle as positive or negative. The results from this initial analysis for each animal sampled are summarised by province in table 4.4. In 2005 68 or approximately 0.7% of samples initially tested positive with the majority in the İstanbul and Tekïrdağ provinces. The number of test positive animals increased dramatically in the Spring 2006 survey (sampled at about the time of the Spring 2006 outbreaks -

see figure 4.1) to 252 or approximately 2% of samples. Again the highest provincial prevalences were in İstanbul and Tekïrdağ but there was also marked increase in the positive samples in Edïrne and Kirklarelï. By Autumn 2006 the number of positive samples was further increased to 358 or nearly 3% of samples. The proportion of test positive samples in Edirne had decreased with the majority of positive samples being from cattle in İstanbul, Kirklarelï and Tekïrdağ. In both 2006 surveys there were no positive samples reported in the 1088 taken altogether in Çanakkale.

Table 4.4: Dichotomous test results for initial CEDITEST (overall and by province for each survey)

2005 survey

	Edïrne	Kirklarelï	Tekïrdağ	Çannakale	İstanbul	Overall
Test positive	6	1	18	2	41	68
Samples	2624	2304	2880	320	1600	9,728
%	0.23	0.04	0.62	0.62	2.56	0.70

2006 Spring survey

	Edïrne	Kirklarelï	Tekïrdağ	Çannakale	İstanbul	Overall
Test positive	39	25	95	0	93	252
Samples	3,008	2,364	3,392	512	3,072	12,348
%	1.30	1.06	2.80	0.00	3.03	2.04

2006 Autumn survey

	Edïrne	Kirklarelï	Tekïrdağ	Çannakale	İstanbul	Overall
Test positive	16	53	158	0	131	358
Samples	3,712	2,688	3,840	576	1,600	12,416
%	0.43	1.97	4.11	0.00	8.19	2.88

As each survey has at least one test positive sample the initial conclusion would be that the results do not provide evidence to demonstrate freedom from foot-and-mouth disease. In the Thrace surveillance system any test positive samples are re-tested with a further CEDITEST NSP kit and subsequently negative samples are classified as negative (not exposed). Additionally any initially positive samples are tested using a liquid phase blocking ELISA test for "O" and "A" structural protein antibodies . Animals with relatively low "O" *and* "A" antibodies are classified as negative (not exposed) on the basis that an exposed animal would have a higher "O" or "A" titre than a vaccinated animal. Whilst these confirmatory test procedures may give a marked increase in the specificity of the surveillance system (by removing falsely positive classified animals) they will also reduce the overall system sensitivity as some animals that were correctly classified as exposed by the first CEDITEST NSP will be erroneously classified as not exposed by the confirmatory tests. Furthermore, as there is sparse evidence describing the overall performance of this confirmatory system use of final classification

for estimation of disease prevalence is difficult. Hence I will confine the remainder of this exploratory analysis to the results from the initial CEDITEST NSP test.

The aggregated provincial data can not be used to draw detailed inferences about the prevalence of exposure to foot-and-mouth disease as it is derived from clustered sampling (at a province and village-level) with differing sample sizes in some villages. Simple estimates of regional prevalence would be biased and would underestimate the uncertainty (confidence interval) around point prevalence estimates. In the next chapter I apply a multilevel mixture model to these surveillance data to estimate the disease prevalence at the different levels of the population.

The dichotomous test results by village are shown for all villages in figure 4.3 and for villages with 64 samples per village in table 4.5. The numerical results for villages with samples sizes different from 64 are shown in table 4.6. These results are suggestive that the prevalence in the samples from each village comes from a zero-inflated process i.e. there are more zero value records than would be expected from common statistical distributions. This is common in infectious disease prevalence data in clustered groups of individuals — the prevalence in groups will be a mixture of zeros for unexposed groups and a distribution or a mixture of distributions reflecting results from exposed groups.

Table 4.5: The number of test positive animals in each village where 64 samples were taken (by survey)

	0	1	2	3	4	5	6	7	8	9	12	13	14	16	19	20	21	26	28	32	42
2005	122	14	4	2	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
2006A	103	38	17	10	9	3	2	1	0	0	0	1	0	0	1	0	0	0	1	0	0
2006B	143	7	10	5	5	2	1	4	2	2	1	1	1	1	1	1	1	1	0	1	1

Table 4.6: The number of test positive animals in each village where other than 64 samples were taken (by survey)

2005		2006A		2006B	
positive	samples	positive	samples	positive	samples
2	63	13	128	3	38
1	39	3	60	0	23
0	13	3	87	0	77
0	14	1	41	0	28
0	6	0	8	0	26
0	4	0	34		
0	11	0	12		
0	8	0	10		
0	8	0	17		
0	65	0	47		
0	11				
0	53				
0	25				
0	38				
0	50				
0	14				
0	26				

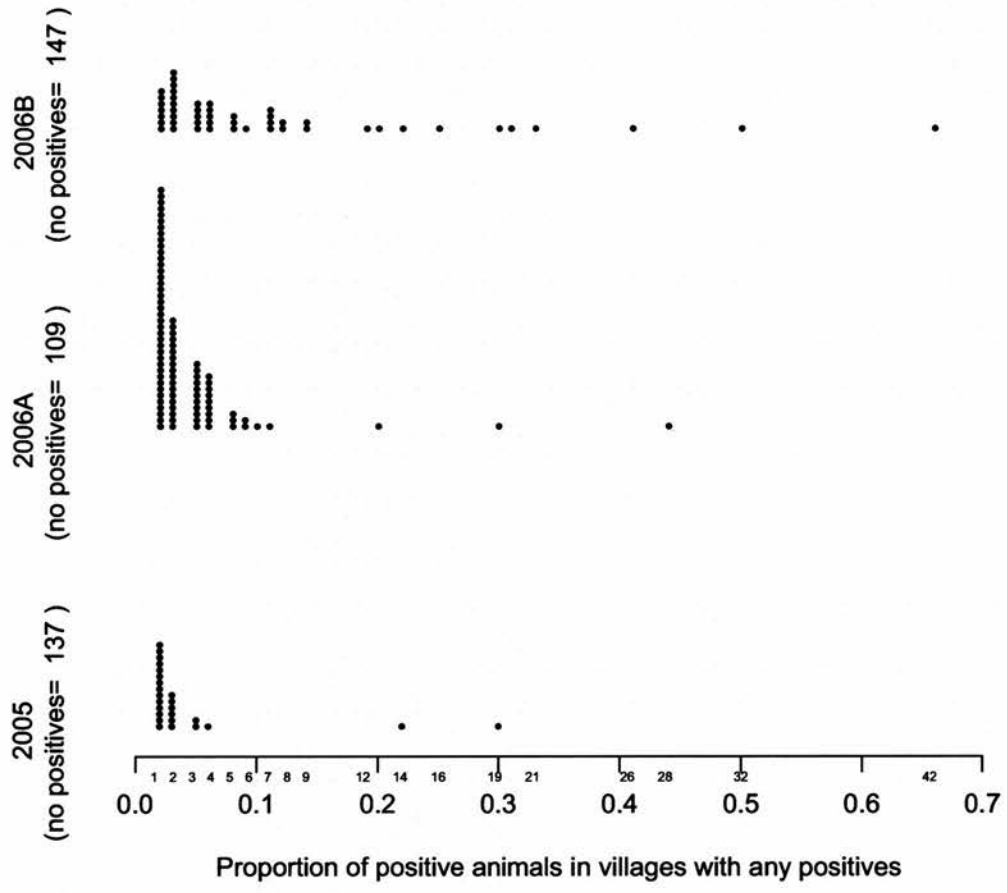


Figure 4.3: The proportion of CEDITEST NSP positive animals in each sampled village by survey. The major x-axis is the proportion with the minor x-axis indicating the number of animals positive to give that proportion if 64 animals were sampled. Each village's result is rounded to two decimal places of proportion and identical rounded results are stacked.

4.4.5 Continuous results from CEDITEST NSP testing

As described in the introduction to this chapter the selected animals in Thrace for sero-surveillance are tested using a CEDITEST NSP test kit. This ELISA based test produces an optical density result which is then standardised by use of a control serum result to give a *percentage inhibition* result. The percentage inhibition result is then used to classify animals as exposed or not-exposed using a simple dichotomous decision rule whereby animals with percentage inhibition results greater than or equal to 50% are classified as exposed. The continuous percentage inhibition results for all sampled animals by each survey are summarised in figure 4.4. The majority of percentage inhibition results are distributed in roughly symmetrical "bell-shaped" distributions with long upper and lower tails. Inspection of this figure suggests a shift in the location of the percentage inhibition results between the Spring 2006 (2006A) and Autumn 2006 (2006B) surveys. This is more easily appreciated by examining an empirical cumulative distribution plot shown in figure 4.5 or a kernel density smoothed plot of the individual results (figure 4.6). The mean percentage inhibition for all animals sampled and tested with the first CEDITEST in 2005 is 9.44 and Spring 2006 is 9.83 as compared to 21.41 in Autumn 2006.

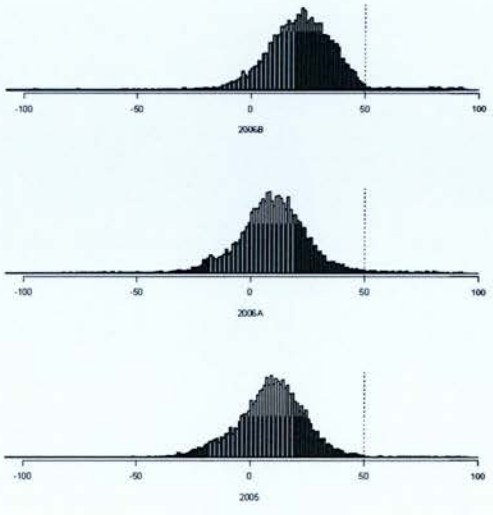


Figure 4.4: Histogram of percentage inhibition results of all sampled animals in Thrace for survey in 2005 & both surveys in 2006. Histogram bins are 1 percentage inhibition unit wide. The vertical dotted line indicated the manufacturers cut-off value of 50 %.

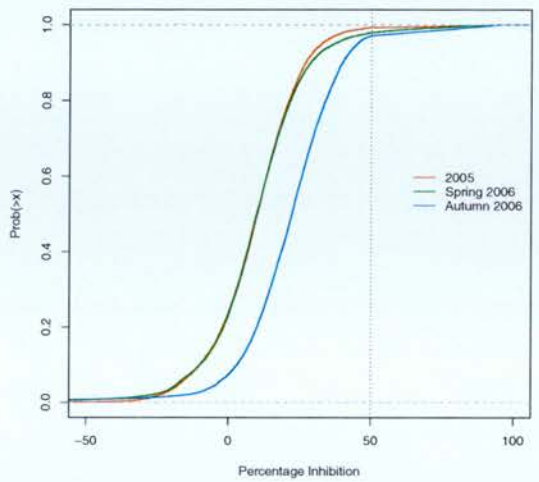


Figure 4.5: Empirical cumulative distribution function of CEDITEST NSP percentage inhibition results by survey. Vertical dotted line at 50 represents manufacturer's cut-off value.

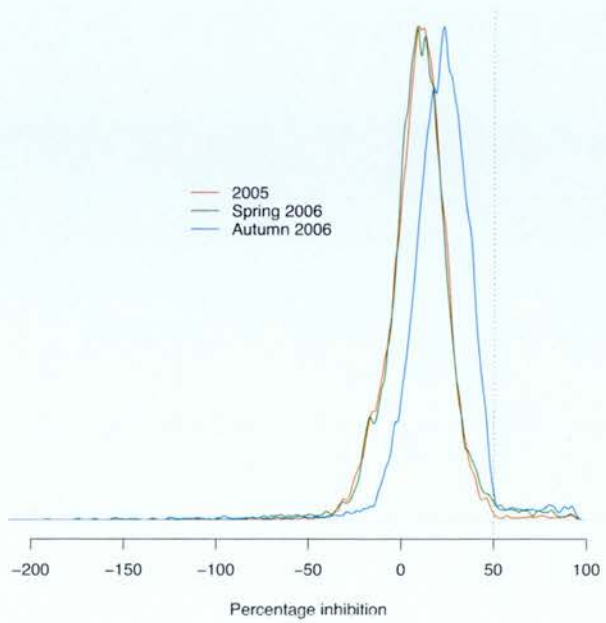


Figure 4.6: Kernel density plot of percentage inhibition results by survey

Figure 4.7: Individual animal percentage inhibition results from the CEDITEST NSP test. The horizontal axis indexes the animals grouping them by survey and within survey animals are grouped by village.

This shift may represent a genuine change in the NSP antibody levels in the sampled livestock or may represent an artefact arising from the analytical methods. Possible analytical causes include:

Control Sera The percentage inhibition result is mathematically derived from the measured optical density of the ELISA test well for the sample scaled by the mean observed optical density of two control wells per ELISA plate (96 well plates). If the control sera reactivity changes between batches the sample results will shift in baseline.

Incubation temperature A change in the processing of the ELISA plates may cause a shift in the resultant optical density readings and if these aren't symmetrically distributed around the control readings could cause a shift in the mean value/location of the results.

ELISA reagents hardware As with incubation temperature

Laboratory technique As above although the CEDITEST NSP appears to be relatively robust to changes in the protocol such as washing techniques [VPH Test standardisation group, DEFRA, 2007].

Only partial data were available describing the raw results used to calculate the percentage inhibition results. For the initial CEDITEST NSP there are 90 control values available (means of paired control sera) for the Spring 2006 survey and 138 values for the Autumn 2006. The surveys require 138 plates suggesting that 48 control sera/plate records are unavailable for the Spring survey. However assuming that the data are missing completely at random these results can be used to estimate any shift in control sera/analysis characteristics. The summary statistics for the available control data are shown in table 4.7 and figure 4.8

Table 4.7: Mean and standard deviation from available CEDITEST control wells pair means for 2006 surveys

	mean	sd	n
2006 Spring	1.28	0.22	90
2006 Autumn	1.25	0.19	138

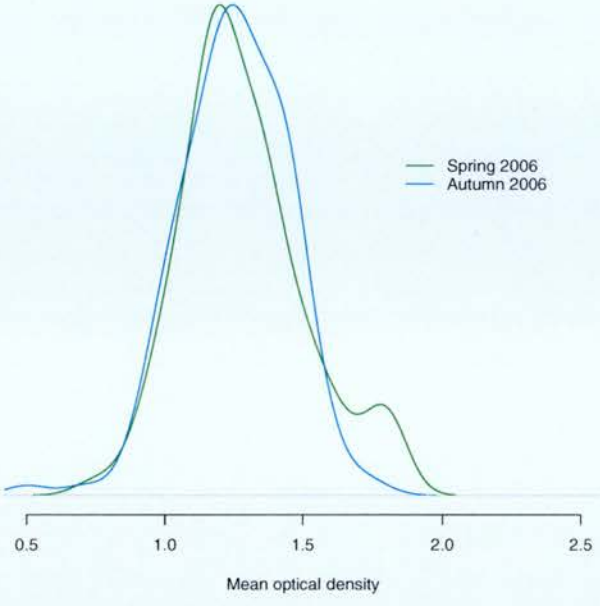


Figure 4.8: A smoothed kernel density plot of the mean paired ELISA control sera results (optical density) for both 2006 surveys

These results show a small (mean 3%) decrease in control well optical density between the Spring and Autumn 2006 surveys. A simulation study shifting the Spring 2006 control values by this magnitude and calculating the resulting percentage inhibition for the samples did not reproduce the observed shift. So although only partial data were available for control sera I consider it unlikely that a change in control sera explains the shift in results between surveys.

No data are available to explore the effects of changes in laboratory technique on the test results although the calculation of percentage inhibition relative to a control sample may negate the effects of gross changes to the reaction or its reading.

Alternatively the observed shift in percentage inhibition results between Spring and Autumn 2006 may represent a shift in the NSP antibody levels in the sampled cattle or another biological factor present in the serum that mimics the NSP antibodies in the ELISA assay. A spontaneous shift in natural proteins is biologically unlikely. However an increase in NSP antibody levels due to increased exposure of the livestock is possible through either increased exposure to foot-and-mouth disease or as a consequence of vaccination although the former is unlikely.

A possible cause of increased exposure of the population would be a shift in the demographic across the surveys with a higher proportion of older cattle in the Autumn 2006 survey. This would give more time for the cumulative effects of repeated vaccination to increase the serum NSP antibody levels in the cattle population. Demographic data were available for the majority of animals sampled although a variety of recording systems were used including estimated date of birth and estimated age. The distribution of cattle ages by survey are shown in figures 4.8 and 4.9.

Table 4.8: Counts of tested livestock classified by age class (months) and survey

Survey	Age	Count
2005	0–6	2701
2006A	0–6	2332
2006B	0–6	2778
2005	6–12	3271
2006A	6–12	5151
2006B	6–12	4771
2005	12+	2892
2006A	12+	3968
2006B	12+	4665
2005	unknown	863
2006A	unknown	898
2006B	unknown	201

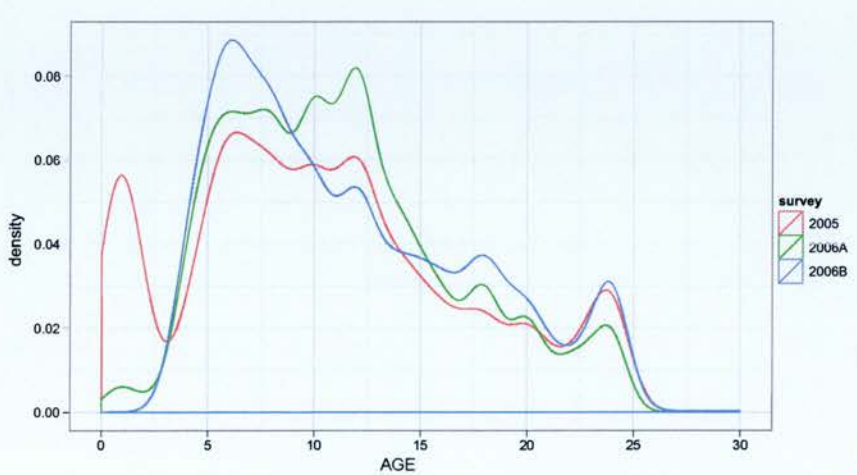


Figure 4.9: The distribution of reported cattle ages (Months) by survey for cattle with a reported age. The data are smoothed with a Gaussian kernel smoother.

The age distributions are broadly similar for the three surveys with the most notable difference being a predominance of cattle aged 0–3 months in the 2005 when compared to the 2006 surveys. The effect of cattle age on the shift of percentage inhibition of the NSP ELISA test is shown in figure 4.10. The Autumn 2006 survey appears to produce consistently higher percentage inhibition values across all age classes suggesting that age effects alone do not explain the shift.

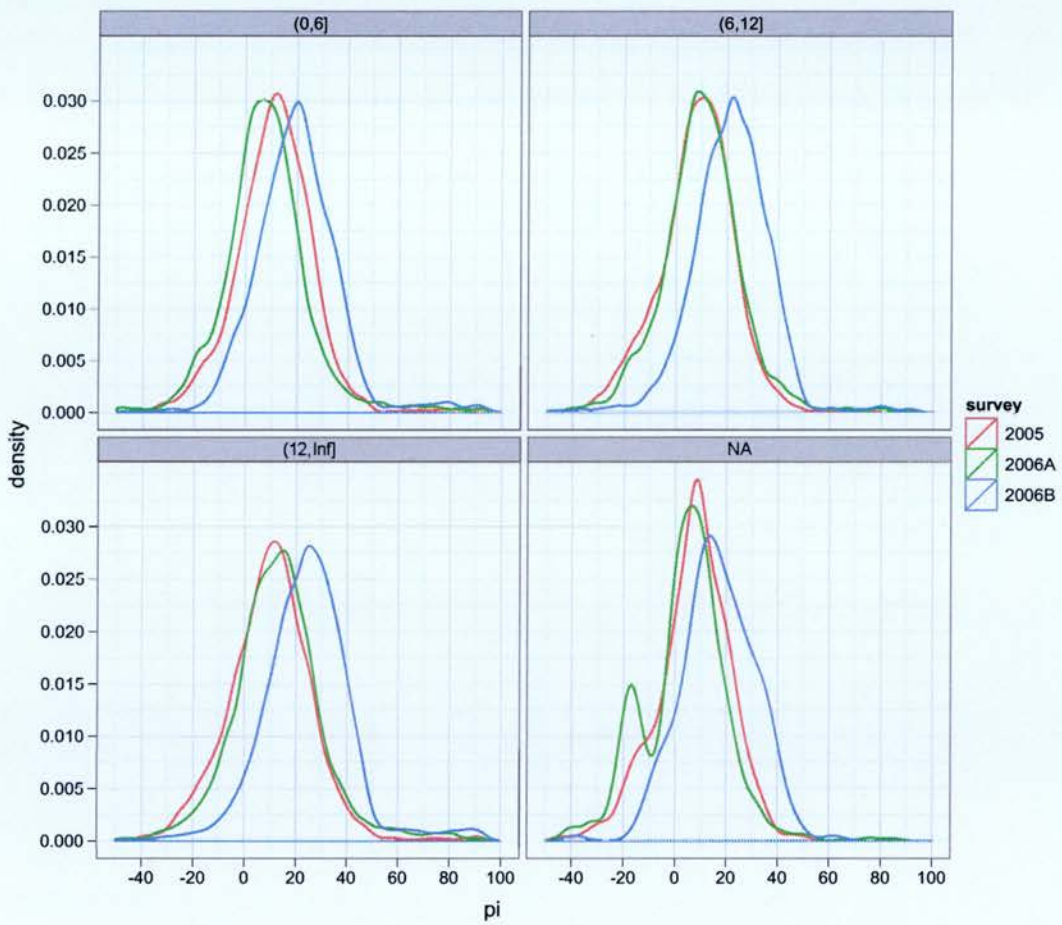


Figure 4.10: The NSP ELISA percentage inhibition results stratified by cattle age class including the cattle with no age records (shown in the 'NA' category). The percentage inhibition results for each class are plotted using a gaussian kernel smoother to show the summary distribution of results.

It is possible that the foot-and-mouth disease vaccinations used in Thrace are not completely free of non structural proteins (EUFMD research group, personal communications 2006). Hence it is possible that the vaccine will potentially elicit an NSP antibody host response (Satya Parida, IAH Pirbright, personal communication 2008).

The shift in percentage inhibition results may have important consequences regarding the interpretation of the test results from Thrace sero-surveillance. For all surveys the manufacturer based cut off value of 50 % was used whereby any animal having a percentage inhibition result greater than or equal to 50 % is classified as exposed. If the shift is common across all animals this will result in animals that would have be classified as unexposed being classified as exposed. Alternatively this could be considered as increasing the sensitivity of the test but dramatically reducing its specificity.

The exact nature of this shift is unknown but by assuming it is an additive effect it is possible to estimate the results that would be seen without the shift.

Visual re-examination of figure 4.6 suggests that rather than a skewing or kurtosis the distribution has shifted largely due to a rightwards translation of the majority of the samples — mathematically each sample's percentage inhibition value is about 12 percentage units higher than it would be without the shift effect. To correct the shift the additive component needs to be estimated and subtracted from the Autumn 2006 results. To use the mean result from each survey to estimate the shift may overly bias the results as it is possible that there has been a true change in sero-prevelance/exposure between the surveys. Hence I have estimated the shift by using the mean of an empirically judged subset of the percentage inhibition results from each survey — the data lying between the 35^{th} and 65^{th} quantiles. This range contains 30% of the data but is unlikely to contain an substantial proportion of exposed animals and hence should be robust to changes in natural exposure of the population. This approach gives a calibration of +0.124 and -12.076 to the Spring 2006 and Autumn 2006 surveys relative to the 2005 survey taking 2005 as a reference state (which is reasonable if the effect is, for example, a consequence of repeat vaccine exposure). The resulting continuous re-calibrated percentage inhibitor results are shown in the density plots in figure 4.11 and the empirical distribution function plots in figure 4.12. Table 4.9 shows the consequent dichotomous results by survey if these values are used with a cut-off of 50 to classify animals exposure status. Three more animals are classed as exposed in 2006 Spring and one hundred less in Autumn 2006 compared to the un-calibrated results.

The effect of the re-calibration of the test results on the Autumn 2006 survey is shown in figure 4.13. Although the number of test positive animals in 30 villages was reduced

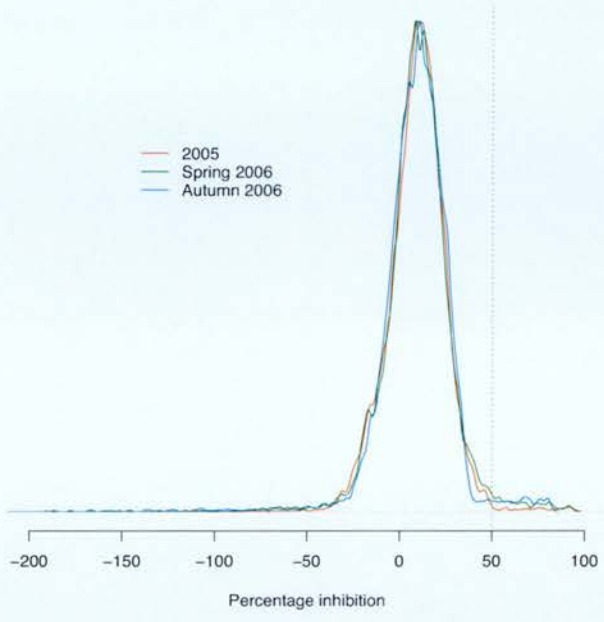


Figure 4.11: Kernel density plot of percentage inhibition results by survey — after CEDITEST re-calibration as described in text

Table 4.9: Dichotomous results of CEDITEST NSP by survey before and after adjustment of percentage inhibition results from Autumn 2006 survey.

	Negative	Positive	Negative (re-calibrated)	Positive(re-calibrated)
2005	9660	68	9660	68
2006 Spring	12096	252	12093	255
2006 Autumn	12058	358	12158	258

only 2 villages were reclassified from having more than one test-positives animals to having no test-positive animals.

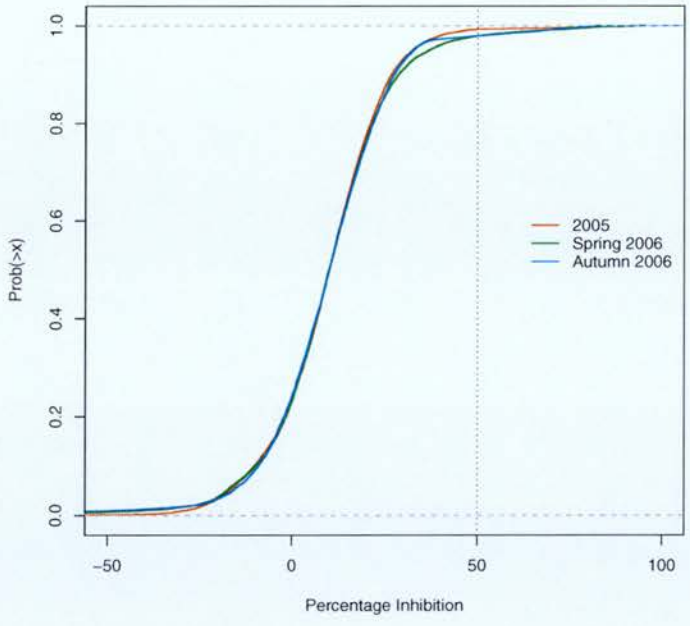


Figure 4.12: Empirical cumulative density plot of percentage inhibition results by survey — after CEDITEST re-calibration as described in text

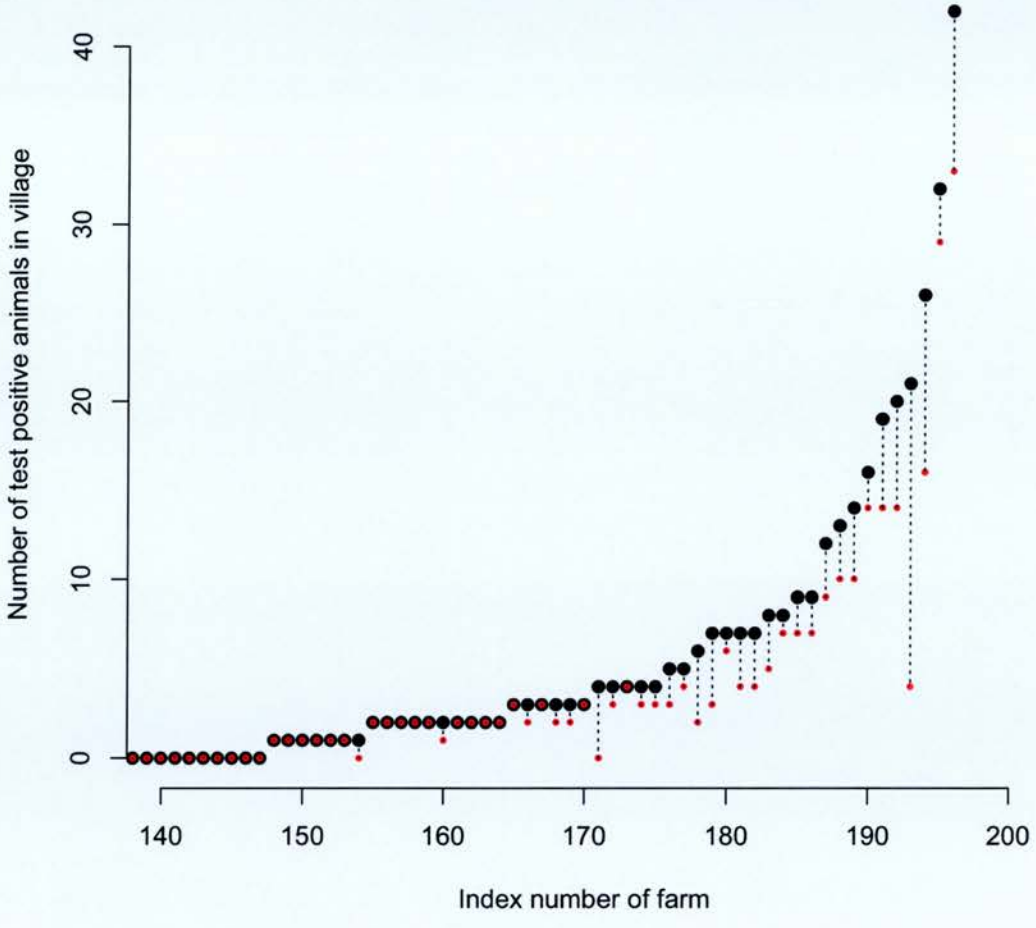


Figure 4.13: The effect of percentage inhibition re-calibration and re classification of CEDITEST results on the Autumn 2006 survey results. The majority of villages have zero test positive animals and are truncated on this figure. The dots show the number of test positive animals in each village before (BLACK dots) and after (RED dots) re-calibration.

4.5 Discussion

Each of the three surveys has produced initial CEDITEST NSP results with a marked number of test positive results suggesting exposure of villages to foot-and-mouth disease. There appears to be a higher level of exposure in the surveys of 2006 compared to 2005. Most villages with test positive results have low rates of positive results but in each survey there are a small number of villages with a markedly high proportion of positive results. These results may represent a background *noise* of false positives due to imperfect test specificity combined with a few truly exposed villages having high rates of positive tests. Alternatively there may be an epidemiology where many villages are exposed but in only a few do outbreaks become large.

The apparent shift in continuous test results between the two surveys of 2006 may have either an analytical or epidemiological explanation. Determination of this would require investigation of laboratory processes and possibly experimental studies or assessment of multiple vaccination effects on diagnostic test specificity. In the absence of these facilities the results may be re-calibrated using the described approach. Although this re-calibration changes the status of approximately 100 animals in the last survey the consequences regarding village status are minor.

As the survey tests results arise from a complex sampling process and are the consequence of analysis with a diagnostic test that will have imperfect sensitivity and specificity simple, unbiased, estimates of village and animal-level prevalence can not be made from the data. In the next chapter I take these data sets and use them to parameterise a multilevel model of village and animal results to estimate the change in epidemiology of foot-and-mouth disease in Thrace over the survey periods.

Chapter 5

Thrace sero-surveillance: Mixture model analysis

5.1 Summary

In all three foot-and-mouth disease surveys over the period 2005–2006 in Thrace a marked proportion of villages contained test positive animals. Classical interpretation of these surveys requires follow-up testing of test positive animals to confirm their serological status. If test-positive animals remain after this follow-up it is considered that there is not evidence to substantiate freedom from disease [OIE, 2005b]. Here I use an alternative multilevel mixture model approach to estimate the underlying epidemiology of disease in the Thrace region over the three surveys and to update estimates of the diagnostic test performance. Previously, zero inflated beta binomial prevalence models have been used for these analyses; with the Thrace data a zero inflated mixture of two beta-binomial prevalence models provides a better fit to the data. Using this model the apparent conclusion is one of high village-level prevalence of disease exposure across the three surveys. Alternatively it may be that disease exposure is relatively low and that the major, low prevalence mixture component, represents animals that are test positive due to antibody responses to the foot-and-mouth disease vaccination programme in Thrace.

5.2 Introduction

Traditionally, disease freedom surveys are designed such that if any animal in the survey tests positive the region should no longer be considered disease-free. In this approach any animals initially testing positive are re-evaluated using confirmatory tests, of high specificity and sensitivity, so that initial false-positive results can be excluded. This interpretation, that one or more confirmed positive animal allows rejection of the hypothesis of disease freedom, derives from the paradigm of traditional survey design where the survey is designed to sample *at least* one exposed herd given a stated herd prevalence with a stated confidence. And sample *at least* one exposed animal with a stated confidence in each sampled herd given a stated animal-level prevalence [Cameron, 1997].

An alternative approach to examining the results of sero-surveillance data is to adopt a Bayesian approach to probability and knowledge by representing our belief about exposure status with a probability using the results from the survey to update our *beliefs* about the region and its animal groups disease status [Suess et al., 2002; Branscum et al., 2004].

This approach allows probabilistic statements to be made about the disease status of the region that give a finer degree of information than the simple disease free or not disease free statements of traditional approaches. In the case of a region, such as Thrace, where it is likely that disease freedom has not been achieved the results of a Bayesian analysis may reveal trends in prevalence and probability of disease freedom. For example it may be possible to demonstrate a decreasing prevalence at a village-level demonstrating an improved disease-control/incursion situation.

In this section the available sero-surveillance data from the 2005 and two 2006 surveys is used to parameterise a Bayesian model describing the epidemiology and diagnostic test performance in the Thrace region. This model provides estimates of test performance which may be used for design and analysis of surveys and estimates of epidemiology which may also be used for survey design and to assess the trend of disease control.

5.3 Methodology

The animal husbandry and disease surveillance structure within Thrace is inherently multi-level hierarchical; cattle exist in groups owned by an individual farmer, these farmers live and manage their cattle within villages and the villages are located within the Thrace region. *A priori* it is reasonable to assume that presence or absence of exposure to disease will be more highly correlated within the village/survey groupings than between them. Hence to obtain unbiased estimates of the prevalence of exposure to disease and to correctly apportion uncertainty in these estimates it is desirable to apply a hierarchical multi-level model for the analysis of the surveillance data-sets [Goldstein and McDonald, 1988].

Conceptually a three level model would appear appropriate to capture the village/owner/animal survey levels of the data. However owners within Thrace villages will tend to graze their animals communally (Nick Honhold, personal communication, May 2007). Consequentially it may be reasonable to assume that a two-level (village/animal) structure is adequate to describe disease in Thrace whereby the probability that an individual animal is exposed is constant across an entire village rather than being clustered by owner/group. This simplification reduces the number of parameters estimated by the model (the hyper-parameters for the distribution of within owner/group prevalence) and further means that the model will only be estimating the latent status of about 150 villages per survey rather than the latent status of many hundreds of owner's herds if a three level model is used. This will increase computation speed and potentially reduce parameter uncertainty if the model is appropriate compared to a three level model.

I apply a Bayesian multi-level model to three sets of sequential surveillance data using dichotomous test results (positive/negative CEDITEST ELISA results). To reduce bias in the estimates the model includes the observational uncertainty by modelling the CEDITEST diagnostic test as imperfect using priors for its sensitivity and specificity.

Data were used from the 2005 and two 2006 sero-surveillance exercises. These data-sets include province, district, village, owner and animal records together with the continuous, percentage inhibition and dichotomised results of the first CEDITEST applied to each animals serum sample. Additionally the data-sets include repeat CEDITEST and liquid-phase blocking ELISA tests for structural protein antigens tests for all sera

positive at the first CEDITEST. For this analysis the first CEDITEST result was used as data are readily available for the diagnostic sensitivity and specificity of this test used singly as opposed to limited data for its use in a serial testing scheme where only performance with regard to carrier detection is available [Paton et al., 2006].

5.3.1 Detailed model structure - base model

Table 5.1: Nomenclature for Bayesian model descriptions and code

pos_i	Number of positive animals in sample from village i
ap_i	Apparent prevalence in sample from village i
n_i	Number of samples taken in village i
se	Sensitivity of diagnostic test
sp	Specificity of diagnostic test
p_i	True prevalence in sample from village i
p_i^*	Prevalence in village i if it is has any exposed animals
$v_{status,i}$	Exposure status of village i (0=unexposed, 1=exposed)
$a_{a,k}$ & $b_{a,k}$	Beta coefficients of animal-level prevalence distribution class k
$class_i$	Prevalence class (1 or 2) of village i
$v_{prev,s}$	Prevalence of village exposure in survey s
$v^*_{prev,s}$	Prevalence of village exposure if the region is exposed
μ_1 & ϕ_1	Mean and variability measures of animal-level prevalence distribution for class 1
μ_2 & ϕ_2	Mean and variability measures of animal-level prevalence distribution for class 2
α & β	Parameters of vague gamma distribution for *phi*
$r_{status,s}$	Region exposure status in survey s
$r_{prob,s}$	Probability that the region is exposed in survey s
μ_v & ϕ_v	Mean and variability measures of village-level prevalence distribution
$p.a.sim.1$	Simulated distribution of within village prevalence for distribution class 1
$p.a.sim.2$	Simulated distribution of within village prevalence for distribution class 2
$v.prev.sim$	Simulated distribution of village-level prevalence

The base model structure uses the hierarchical model of Branscum et al. [2006] which uses a two-level (group and individual) structure: in this case villages and animals. The observed data are the counts of test positive animals in the sample from a given village. These are conditional on the unobserved prevalence of exposure in each village and the diagnostic performance of the CEDITEST. The count of test positive animals in each village is modelled as:

$$pos_i \sim binom\left[n_i, p_i \times se + (1 - p_i) \times (1 - sp)\right]$$

Where pos_i is the number of test positive animals in village i, n_i is the sample size in the village i, se is the CEDITEST sensitivity, sp is the CEDITEST specificity and p_i is the prevalence of exposed animals in the sample from village i.

The analysis makes the assumption that the number of test positives in the sample is the result of a simple binomial process, i.e. the sampling and testing is performed *with replacement* [Cameron, 1997]. In reality this is not the case. To accurately capture the sampling mechanism requires an intermediate modelling step using a hypergeometric distribution whereby a village containing animals exposed with a prevalence p_i is sampled without replacement to generate a sample with a number of exposed animals distributed hypergeometrically. Unfortunately it is not technically possible to include this form of model and whilst modelling the diagnostic test as imperfect within the WinBUGS /JAGS [Plummer, 2003] environment used to analyse these data. The hypergeometric step is only important when sample sizes are large relative to village sizes - in this case although the village sizes for each sample are not recorded inspection of the demographic data (2005 sample frame) and manual matching suggest that the villages are generally much larger than the modal sample size of 64 (see chapter 4 section 4.4.3).

The true prevalence in each village is not constant across all of the villages in the survey. It will be zero in unexposed villages and variable (due to the dynamics of infection and variability in biology/husbandry) in exposed villages. In the base model p_i, the prevalence in village i is modelled as a mixture distribution of zeros (for non-exposed villages) and a beta distribution (for exposed villages). The beta-binomial distribution attempts to capture the over dispersion (above a binomial) in prevalence seen in exposed villages [Johnson et al., 2004; Cameron and Baldock, 1998b].

$$\begin{array}{lcll} p_i & \sim & beta(a_{a,1}, b_{a,1}) & \text{village exposed} \\ & = & 0 & \text{village NOT exposed} \end{array}$$

The status of a village, $v_{status,i}$ of village i as exposed or not exposed is modelled as a bernoulli variable:

$v_{status,i} \sim bern(v_{prev,s})$

Where $v_{prev,s}$ is the prevalence of exposed villages (in survey s).

The prevalence of village exposure (i.e. the probability that a randomly selected village is exposed) given that the region is exposed is not fixed from survey to survey and is modelled as a beta distribution:

$v_{prev,s} \sim beta(v.prev.a, v.prev.b)$

Where $v.prev.a$ and $v.prev.b$ describe the distribution of village-level prevalence across surveys. If the region is not exposed the village prevalence $v_{prev,s}$ is zero.

The regional status (exposed or not exposed) for survey s, $r_{status,s}$, is modelled as a bernoulli variable:

$r_{status,s} \sim bern(r_{prob,s})$

Where $r_{prob,s}$ is the probability that the region is exposed at the time of survey s.

5.3.2 Model modification — three distribution mixture

The base model, as derived from Branscum et al. [2006] was modified to attempt to capture the apparent heterogeneity in within village prevalence observed in the exploratory data analysis 4.4.4. Examination of the prevalence data suggested that the prevalence in exposed villages may be *over-dispersed* beyond even a beta-binomial distribution and was bi-modal. Hence in this chapter I have redefined the Bayesian model to use a mixture of three distributions for within village prevalence:

$$\begin{array}{lcll} v_{prev} & \sim & beta(a_{a,1}, b_a) & \text{village exposed and low prevalence} \\ v_{prev} & \sim & beta(a_{a,2}, b_a) & \text{village exposed and high prevalence} \\ & = & 0 & \text{village NOT exposed} \end{array}$$

Where $a_{a,1}$ $b_{a,1}$ and $a_{a,2}$ $b_{a,2}$ are the beta parameters of the low and high prevalence villages respectively. The parameter $p.2$ represents the probability that an exposed village has a prevalence drawn from the high rather than low prevalence group. In the JAGS model code the mean village prevalence for the high group is modelled as the mean from the low group with an additional scaling increment to ensure identifiability of the parameters Stephens [2000].

5.3.3 Data requirements

The model uses the number of test positive animals and the sample size for each village sampled in each of the three surveys. The data were re-structured for analysis using Microsoft Access (Redmond, USA) and R [R Development Core Team, 2007]. The models were run with both the original data and also with 'corrected' data where the

continuous CEDITEST results were shifted and reclassified to correct for the apparent change in percentage inhibition values observed between the two surveys in 2006 as discussed in chapter 4.

5.3.4 Choice of prior distributions

As the model is fully Bayesian it is necessary to provide prior distributions for all parameters in the model. Part of the development and evaluation of such models is assessing the results conditional on different prior distributions so for some parameters a number of different prior distributions were tested.

The diagnostic test sensitivity and specificity, *se* and *sp* were assigned beta distribution priors using diagnostic test evaluation data from Engel et al. [2008] derived from Brocchi et al. [2006]. These results estimate the sensitivity and specificity of the diagnostic tests in vaccinated animals as the majority of cattle in Thrace are vaccinated (based on preliminary vaccination coverage data examined at EUFMD research group meeting May 2006). The Bayesian analysis from [Engel et al., 2008] removes the assumption that a perfect gold standard is available to validate the estimation and provides an uncertainty distribution for the sensitivity and specificity (described for efficiency with a beta distribution using the parameters estimated for the stochastic simulation model in Chapter 3). As these sensitivity and specificity priors were relatively precise the model's sensitivity to prior information was assessed using less informative 'vague' priors. The approximate form of these distributions is shown in figure 5.1. With a Bayesian analysis it is critical that the prior distributions are not informed by the data. As there were no available studies estimating prevalence distributions in the villages and animals (other than expert opinion that would inevitably be informed by the survey results) vague priors were used for animal and village prevalence distributions. The parameterisation derived from Branscum et al. [2004] was used where the two beta distribution parameters for each prevalence distribution were represented thus:

$$
\begin{aligned}
prevalence &\sim \text{Beta}(a, b) \\
a &= \mu\phi \\
b &= \phi(1-\mu) \\
\mu &\sim Beta(0.001, 0.999) \\
\phi &\sim Gamma(1, 0.001)
\end{aligned}
$$

Where μ is the mean of the prevalence distribution with a hyper-prior distribution between almost zero and one (actual limits just within this improve numerical stability). And ϕ is related to the variability of the prevalence and has a Gamma distribution hyper-prior with a mean of 100 and a standard deviation of 100.

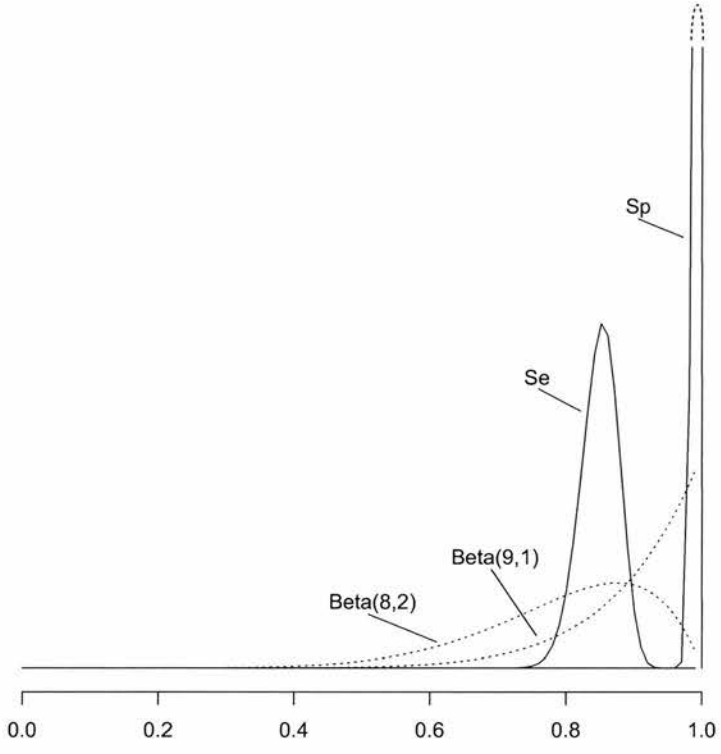


Figure 5.1: Density of prior distributions for diagnostic test performance: 'Se' and 'Sp' from Bayesian estimation of NSP performance by [Engel et al., 2008] and two theoretical, less informative distributions used for both Se and Sp as vague priors. (The prior 'Sp' from is highly informative and the plot has been truncated in the y-axis on this plot)

5.3.5 Model estimation

The model was estimated using Bayesian MCMC with JAGS [Plummer, 2003] software called from R [R Development Core Team, 2007] using the Runjags package [Denwood, 2008]. The JAGS software uses a similar model description language to the more commonly used WinBUGS software Spiegelhalter et al. [1999] but gave faster and more

stable results on initial testing. An MCMC burn-in period of at least 100,000 iterations was used followed by a capture period of at least 100,000 iterations. In cases of poor convergence the burn-in and capture periods were increased up to 500,000 each. Model runs took approximately 6 hours for a total of 200,000 iterations. Results were tabulated and graphed using the R statistical system [R Development Core Team, 2007]. Each JAGS simulation used three chains with highly dispersed starting points for the MCMC chain (i.e. 0.1, 0.5 and 0.9 for diagnostic test parameters and prevalences). Convergence of the MCMC process was assessed by use of the Gelman-Rubin convergence statistic, as modified by Brooks and Gelman [1998]. Model selection between the zero/single beta binomial mixture model and zero/double beta binomial mixture model was primarily based on the deviance information criteria [Spiegelhalter et al., 2002; Celeux et al., 2006]. This is a measure of model fit based on the likelihood integrated over the posterior parameter density with an added penalty derived from the number of parameters estimated by the model to adjust for the normal improved fit of less parsimonious models [MRC Cambridge]. Model selection was further informed by the convergence behaviour of the models assessed by examination of the trace plot of the samples from the posterior distributions and calculation of the Gelman statistic [Brooks and Gelman, 1998] calculated using the R library 'coda' [Plummer et al., 2006]. This is a measure of posterior sample chain stability and provides an estimate of how well the parameter estimates from the posterior samples will capture the *true* uncertainty in the parameters [Cowles and Carlin, 1996].

5.3.6 Model Diagnostics

For a given model, data-set and prior parameter distributions the Bayesian approach provides posterior parameter distributions that precisely reflect how the evidence from the data and model should update our beliefs about the parameter values Lee [1997]. However it is frequently impossible to determine an analytical form for the posterior distributions in anything other than a simple model. Hence it is necessary to use numerical methods to obtain samples from the posterior distributions and hence derive statistic and probability plots. In the WinBUGS / JAGS software these samples are draw using Monte Carlo markov-chain (MCMC) methods. As the number of samples drawn tends towards infinity the distribution of the samples will tend towards that of the posterior distribution of the parameters. Practical speed with which simulations

approach a close approximation to the posterior will vary from model to model and with data and priors. As sample sizes are limited it is necessary to perform a number of diagnostic procedures to demonstrate that the posterior samples used do not have any properties that suggest their distribution does not model the Bayesian posteriors [Raftery and Lewis, 1992]. These methods include use of three or more simulation chains with different starting values for each parameter; visual examination of the simulation history for each simulation chain for each parameter; examination of autocorrelation of each chain; examination of the form of each parameter's posterior distribution and application of formal convergence and effective size statistics. The individual methods will be discussed with reference to the model output in the results section.

5.4 Results

5.4.1 Model selection

The Bayesian model was estimated using both a single and double beta-binomial mixture model with raw and re-calibrated data. Priors derived from Engel et al. [2008] and less informative priors were used. The models that successfully converged on a relatively stable posterior MCMC series are listed in table 5.2. The models are compared by deviance information criteria (DIC) in the table. For a given prior distribution and data set the zero + double beta-binomial models consistently had a lower DIC suggesting a 'better' model fit. With mixture models the number of effective parameters used in the calculation of the DIC is not rigorously defined and hence there are issues regarding the use of DIC for model selection [Celeux et al., 2006]. However the double mixture model in this case had a DIC markedly lower than that for the single mixture models so small effects of DIC estimation methodology were considered unlikely to be significant. For further discussion here the zero/double model based on Engel et al. [2008]'s estimation of the CEDITEST diagnostic test sensitivity and specificity applied to the re-calibrated data set is selected as it demonstrated good convergence, a low DIC and uses literature informed prior information for the CEDITEST. This model's DIC was only marginally higher than a similar model using vague priors and produced similar posterior estimates.

model	data	se_a	se_b	sp_a	sp_b	DIC
single	re-calibrated	141.30	25.30	593.90	6.60	1019.05
single	raw	9.00	1.00	9.00	1.00	1025.19
single	raw	6.00	2.00	6.00	2.00	1027.47
double	raw	141.30	25.30	593.90	6.60	1015.50
double	**re-calibrated**	**141.30**	**25.30**	**593.90**	**6.60**	**998.46**
double	raw	8.00	2.00	8.00	2.00	1015.72
double	re-calibrated	8.00	2.00	8.00	2.00	997.95

Table 5.2: Summary of model runs that ultimately showed stable posterior simulation histories. Deviance information criteria (calculated using pD method Plummer [2003]). The model selected for discussion is shown in **bold** font. se_a & se_b, & sp_a & sp_b are the pairs of parameters of the sensitivity and specificity beta distribution priors describing the performance of the CEDITEST NSP diagnostic test.

5.4.2 Model Diagnostics

As the posterior distribution statistics in the JAGS software are calculated from MCMC samples it is necessary to confirm that the MCMC simulations have adequately converged onto stable time series Plummer et al. [2006]. Initial visual inspection of the selected model trace output (see figure 5.9) showed apparently stationary MCMC chains with no apparent time related drift or divergence.

Visual inspection of the smoothed posterior density plots for the three simulation chains (each having different starting points) showed reasonable agreement between the form of each chain (see figures 5.3 and 5.4) and with the majority of parameters smooth, uni-modal posterior distributions. The exceptions are $p.a.sim.2$ and $mu.2$ representing the mean and distribution of the animal-level prevalence in high prevalence exposed villages. This bi-modality will be considered in the discussion section.

The Gelman statistic [Brooks and Gelman, 1998] for each parameter is shown in table 5.3 as calculated in R using the coda package [Plummer et al., 2006]. A potential scale reduction factor (PSRF) is estimated using the between and within chain variance estimates for chains with over-dispersed starting points. A converged model will have similar variance within a chain as between values in different chains. The resulting PSRF value for each parameter represents an estimate of degree of posterior variance inflation arising from non-convergence — the posterior estimates will have a higher variance than if the model had converged and the chains were mixing. A converged set

of chains should have a PSRF close to unity. In this model each parameter's PSRF is within 5% of 1.0 providing further reassurance that the MCMC chains have adequately converged.

Parameter	Point est.	97.5% quantile
mu.1	1.0004	1.0015
mu.2	1.0004	1.0009
p.2	1.0023	1.0043
p.a.sim.1	1.0020	1.0049
p.a.sim.2	1.0000	1.0000
phi.1	1.0016	1.0030
phi.2	1.0237	1.0244
se	1.0000	1.0002
sp	1.0001	1.0005
v.prev[1]	1.0007	1.0020
v.prev[2]	1.0009	1.0020
v.prev[3]	1.0010	1.0026
v.prev.sim	1.0000	1.0001

Table 5.3: Gelman convergence diagnostics for zero/double model fitted to re-calibrated surveillance data using *se* and *sp* priors from [Engel et al., 2008]

The R coda package also provides a method by which the effective size of a sample set may be estimated [Plummer et al., 2006] based on spectral analysis of the posterior samples Heidelberger and Welch [1981]. This method estimates the number of independent samples that would have the same information content as the (possibly autocorrelated) full posterior sample set. A large reduction would suggest high autocorrelation and redundancy. If the results set is small it is suggestive that a larger set of simulations is required. The effective sample sizes are shown in table 5.4. Even the smallest effective sample size of 927 (for parameter $p.2$) will provide a useful point and distribution estimate. This was confirmed using simulated sampling from a variety of plausible distributions for parameters with a sample size set to 927.

5.4.3 Results from posterior distribution samples

Table 5.5 and figures 5.3 and 5.4 show the results of the posterior distribution samples from the selected model. For all three surveys the *r.status* parameter which captures the estimated exposure status of the Thrace region on each model iteration consistently

Parameter	Effective sample size
mu.1	3910.50
mu.2	1807.48
p.2	927.45
p.a.sim.1	12946.61
p.a.sim.2	4677.02
phi.1	3092.85
phi.2	4011.30
se	14023.93
sp	7604.07
v.prev[1]	5265.45
v.prev[2]	2336.87
v.prev[3]	2508.71
v.prev.sim	11464.64

Table 5.4: Effective simulation set sizes calculated in R using empheffectiveSize see Heidelberger and Welch [1981]

classified the region as exposed with no uncertainty. The *v.prev.sim* parameter represents the distribution of village-level prevalence if the area is exposed i.e. if, in a given survey, the region is exposed how probable is a given village-level prevalence. This distribution has a mean of approximately 0.3 with a wide variation — the model and data suggest that if the Thrace area is exposed the village-level prevalence will be variable from survey to survey. For each survey the model estimates the specific village-level prevalence (parameter *v.prev*[1..3]). These estimates all have a large uncertainty but the mean increases from 0.12 in 2005 to 0.46 in Spring 2006 then drops to 0.28 in Autumn 2006.

Parameters *pa.sim*.1 and *pa.sim*.2 represent the distribution of animal-level prevalence found in low and high prevalence exposed villages. Parameter *p*.2 is the probability that an exposed village is in the *high* prevalence class (2). *mu* and *phi* are hyper-parameters that define *pa.sim*1 and *pa.sim*.2. The results suggest that about 16% of exposed villages are in the high prevalence class. The low prevalence class has a mean of approximately 0.04 with some uncertainty although likely to be less than 0.1. The prevalence in the high prevalence class is less well defined with a mean of 0.24 and a credible interval from 0 to approximately 0.75.

The CEDITEST diagnostic sensitivity and specificity in this model were given informative priors. These priors are shown plotted against the estimated posterior distributions

Parameter	Mean	SD	2.5%	97.5%
mu.1	0.037	0.006	0.025	0.049
mu.2	0.242	0.091	0.069	0.403
p.2	0.167	0.107	0.048	0.465
p.a.sim.1	0.037	0.021	0.006	0.086
p.a.sim.2	0.238	0.166	0.001	0.599
phi.1	140.484	97.303	30.005	396.264
phi.2	11.513	12.737	2.347	40.684
r.status[1]	1.000	0.000		
r.status[2]	1.000	0.000		
r.status[3]	1.000	0.000		
se	0.8467	0.028	0.788	0.897
sp	0.999	0.000	0.998	0.999
v.prev[1]	0.124	0.046	0.052	0.229
v.prev[2]	0.458	0.090	0.310	0.665
v.prev[3]	0.276	0.054	0.190	0.406
v.prev.sim	0.298	0.180	0.024	0.749

Table 5.5: Summary of posterior distributions for parameters of zero/double model fitted to re-calibrated surveillance data using priors from [Engel et al., 2008]

in figure 5.2. The results suggest that the model and data do not add any information to the prior estimates of the diagnostic test sensitivity. However the location and uncertainty regarding the diagnostic test specificity is changed with a mean of 98.9% to a mean of 99.8%.

It is likely that the parameter estimates from this model will be correlated as, for example, a particular apparent prevalence in the data may be explained by high true prevalence and low test sensitivity or the converse. Figure 5.5 shows the approximate degree of parameter cross correlation. The highest correlations are between $p.2$ (the probability that an exposed village has a prevalence in the high prevalence group) and $mu.2$ the mean prevalence of this group; and between $v.prev[2]$ (the prevalence at village-level in Spring 2006) and $mu.1$ the mean prevalence of the low prevalence group of villages. The form of these correlations are shown in figures 5.6 and 5.7. Both pairs of parameters are negatively correlated such that, for example, a high estimate of prevalence in the higher prevalence villages is associated with a lower probability of an exposed village being in this group. The correlation of CEDITEST sensitivity and specificity estimates is shown in figure 5.8.

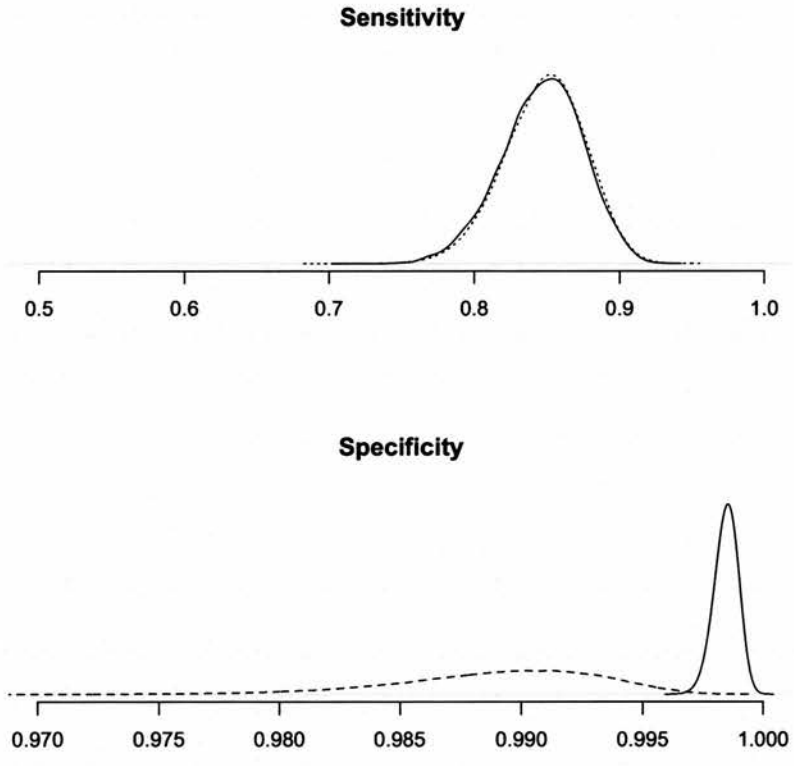


Figure 5.2: Density plots of the prior(shown with a dotted line) and kernel smoothed posterior(shown with a solid line) distributions for the CEDITEST sensitivity and specificity.

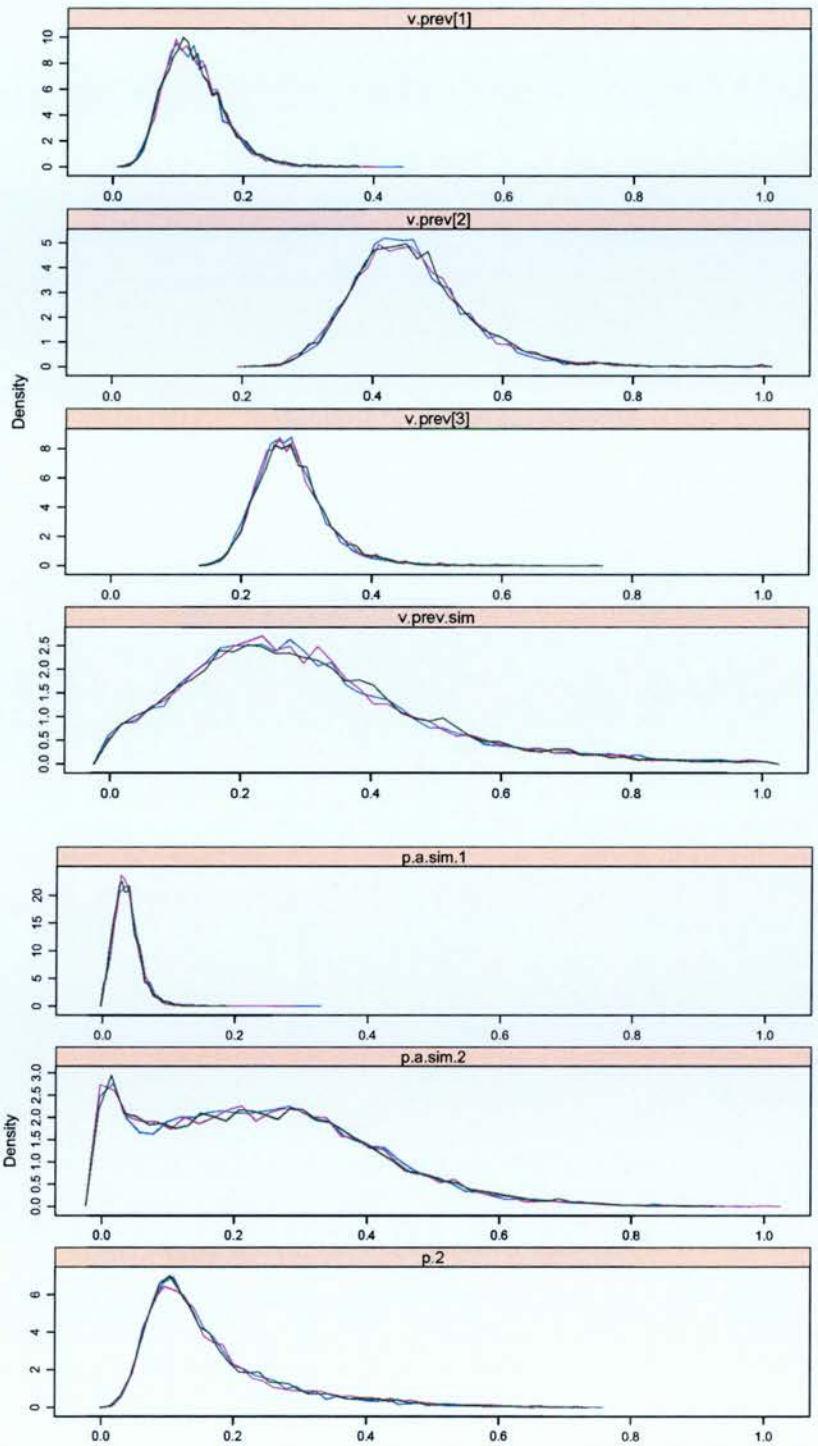


Figure 5.3: Posterior density plot of prevalences using zero/double mixture model with re-calibrated data — priors from Engel et al. [2008] for sensitivity and specificity — coloured lines in each panel show results of each of the three MCMC simulation chains (smoothed using a gaussian kernel window)

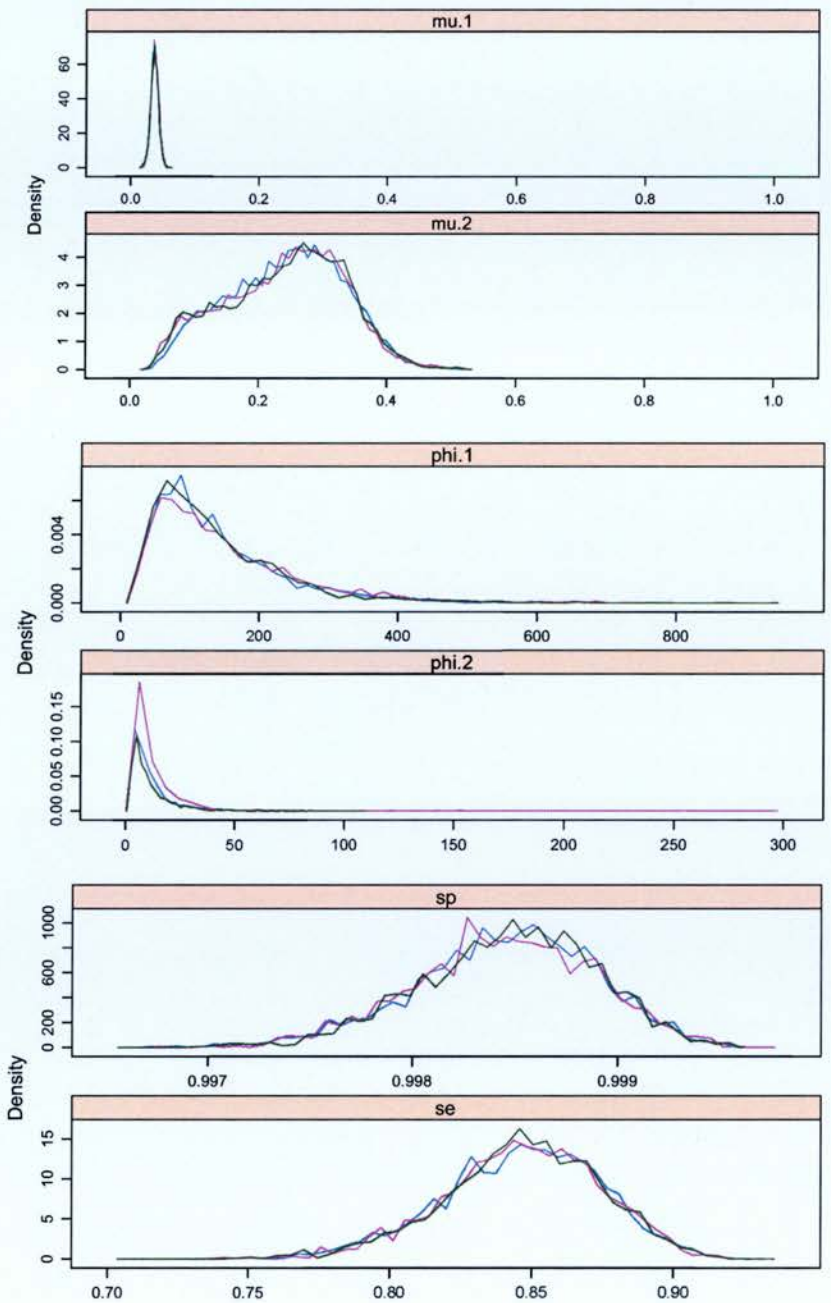


Figure 5.4: Posterior density plot of other parameters using zero/double mixture model with re-calibrated data — priors from Engel et al. [2008] for sensitivity and specificity — coloured lines in each panel show results of each of the three MCMC simulation chains (smoothed using a gaussian kernel window)

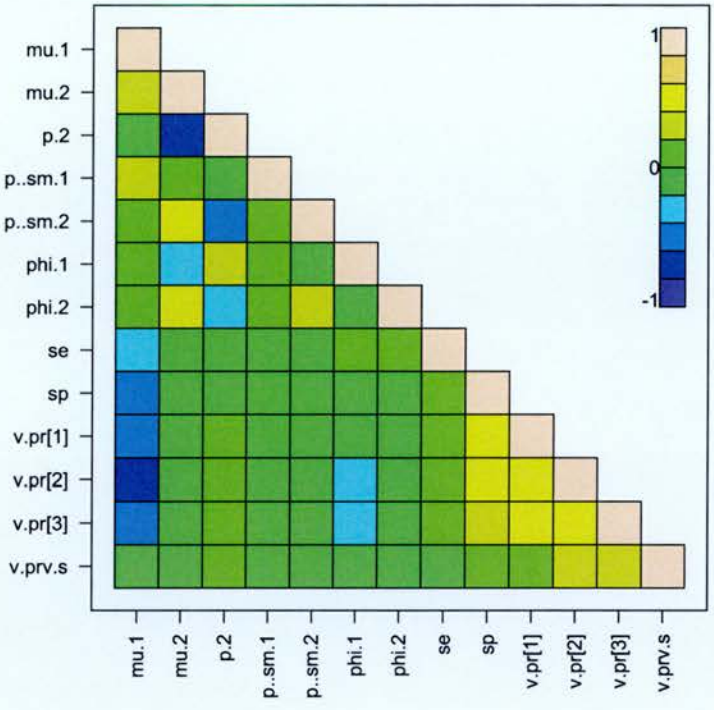


Figure 5.5: Cross correlation of parameter posterior distributions — zero/double mixture model with re-calibrated data using priors from Engel et al. [2008] for sensitivity and specificity

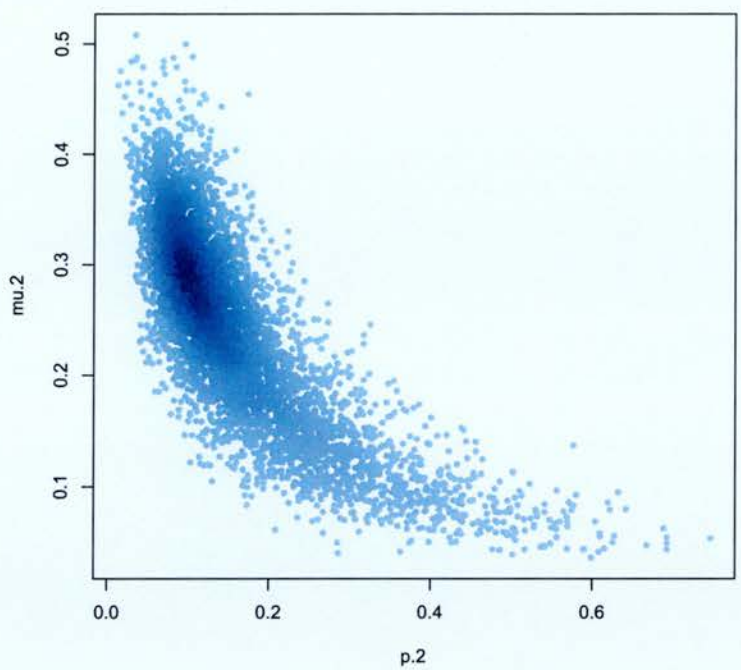


Figure 5.6: Correlation plot of p.2 (probability that an exposed village is from the high prevalence group) against mu.2 (the mean prevalence of the high prevalence group)

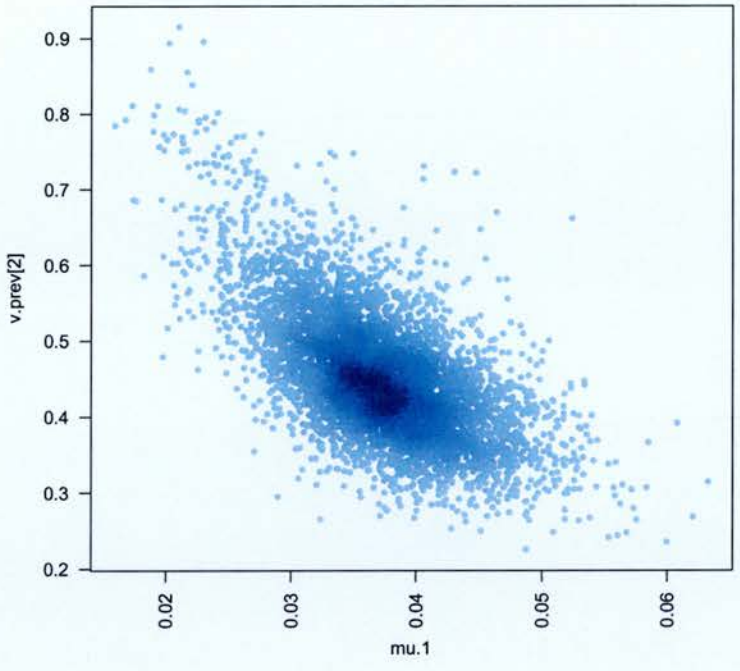


Figure 5.7: Correlation plot of v.prev[2] (the village-level prevalence in Spring 2006) against mu.1 (the mean prevalence in the low prevalence group of villages)

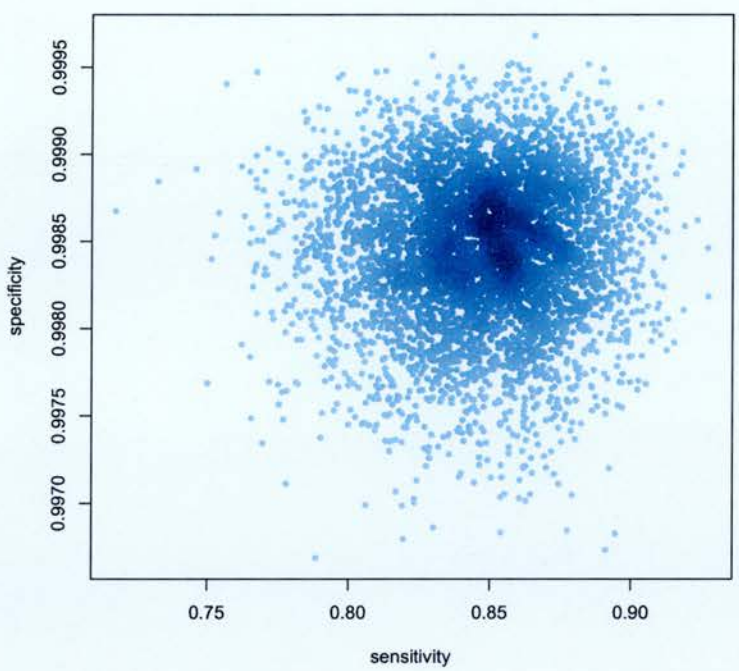


Figure 5.8: Correlation plot of estimated sensitivity of CEDITEST against against estimated specificity of CEDITEST

5.5 Discussion

This analysis classified the Thrace region as exposed at all three surveys given the model, the prior assumptions and the data used to parameterise the model. This result is compatible with ad hoc analysis of the survey data — even in the survey with the fewest test positive animals (2005 survey) there were two villages with high positive counts of 14 and 19 test positive animals out of 64 sampled. Even with a relatively non-specific test (of, for illustration, only 90 or 99 % specificity) to observe this many or more false positive test results in an unexposed sample of 64 animals is extremely unlikely — see table 5.6.

	Specificity = 0.90	Specificity = 0.99
$P(x >= 14)$	0.00136	$1.01 \times e^{-16}$
$P(x >= 19)$	$2.5 \times e^{-06}$	$1.29 \times e^{-24}$

Table 5.6: The probability of observing at least 14 or 19 false positive test results in a sample of 64 animals for two different test specificity's (calculated as $Prob(x < X)$ where $X \sim Binom(64, 1 - Sp)$) with x false positives and Sp the test specificity)

The model also estimates the posterior distribution of village-level prevalence of exposure for each survey. This is the probability that each individual village is exposed in each survey given that the region is infected. With the time profile of NSP antibodies this is equivalent to infection and serological response within the previous 6–12 months. Over the three surveys the village exposure probability increased from 0.12 in the 2005 survey to 0.46 in the Spring 2006 survey (at the time of the Thrace multiple outbreaks) and then reduced to 0.28 in the Autumn 2006 Survey. The trend of these results are compatible with the qualitative observations of the time considering the likely persistence of NSP antibody seroconversion and given that the target age for sampled animals was 4 months to 2 years. This result should be critically interpreted as the probability that a village sample is drawn from either of the two beta-binomial distributions of sample group prevalence (as opposed to being a zero-prevalence village). As there is a non-zero probability that the beta-binomial distributions have a zero prevalence these village-level prevalence results will over-estimate the actual village-level prevalence. Simulated results suggest that with the animal-level prevalences in question this error is of the order of 10%. I have explored models that required an exposed village

to have at least one exposed animal and rigorously replicate the sampling process. Unfortunately these models were not computationally tractable. Even with the correction for the beta-binomial modelling assumption the resulting village-level prevalences are higher than expected from clinical reporting of observed foot-and-mouth disease cases (none in 2005 and the outbreaks in Spring/Summer 2006). There are several possible explanations for the discrepancy:

The posterior estimates of the three village prevalences may be correctly estimating the village-level exposures. The distributions are very uncertain so estimate relatively low village-level prevalences each year with finite probability. As the Thrace population is vaccinated it is possible that cattle movements and trading could result in a high proportion of villages having a low animal-level prevalence of exposed animals — vaccination preventing disease from taking off to create numerous outbreaks.

Alternatively the model may be correctly estimating the expected test status in the Thrace villages but this may not reflect the animal-level exposure and hence may over-represent the village-level prevalence of *true* foot-and-mouth disease exposure. I propose two possible mechanisms for this:

The CEDITEST diagnostic test may not be as specific as estimated in this analysis (and by previous evaluations). Hence false positive test results will be classified as truly positive and result in mis-classification of animals and villages thereby inflating the village-level prevalence. However, the model estimated the test to be very specific (and this specificity estimate was robust to changes in the diagnostic test specificity priors from the informative priors from Engel et al. [2008] to vague Beta(8,2) priors).

Alternatively: The cattle in Thrace are vaccinated using a trivalent foot-and-mouth disease vaccine. It is possible that this vaccine may contain traces of 3 ABC non-structural protein (NSP) which would result in an immune response creating NSP antibodies as detected by the CEDITEST ELISA (Satya Parida (IAH), personal communication 2008]. This would mean that the survey is, in many cases, estimating high levels of village prevalence of NSP antibodies as a consequence of vaccination artefacts rather than as a result of foot-and-mouth disease exposure. This hypothesis would require vaccination coverage or response to be highly variable between villages. Unfortunately there are only limited data available to estimate coverage. Current coverage estimates are based on the number of vaccination doses supplied and demography data from livestock censuses. These estimate coverage at between 80% and 123%. The estimates in

excess of 100% are likely to arise from differences between cattle presented/recorded at census and cattle presented at vaccination although clerical errors are also possible. Investigation of the vaccination-related hypothesis would initially require comparison of animal-level test results with animal-level vaccination results to look for an association. If this was indicated then estimation of the effect may permit some correction. Current surveys are targeted at animals between 4 months old and 24 months old to reduce the effects of repeated vaccination. If sufficient young cattle numbers are available it may be appropriate in future surveys to reduce the upper age limit to 12-18 months, although this will then reduce the overall sensitivity of the system as cattle on average will be less likely to be present during a previous village outbreak.

5.5.1 Mixture and animal-level prevalence parameters

Parameter $p.2$ in the model — the probability that an exposed village was in the higher prevalence group — was estimated with high uncertainty. This parameter was also negatively associated with another uncertain parameter, $mu.2$, the mean prevalence in the high prevalence villages. This is likely to be a consequence of the low numbers of high prevalence results and the inherent problems of identifying parameters of mixture models particularly when one distribution of the mixture has a much lower membership that the other distribution. The information required to parameterise the smaller distribution is effectively reduced by noise in the tail of the larger distribution that overlies it.

The interpretation of the estimates for parameters $p.2$ (defined above), $pa.sim.1$ (the animal-level prevalence distribution in low prevalence villages) and $pa.sim.2$ (the animal-level prevalence distribution in high prevalence villages) is considered jointly. The selected model estimates each exposed village as having an animal-level prevalence coming from one of two distributions ($pa.sim.1$ or $pa.sim.2$) with conditional probability of coming from $pa.sim2$ being $p.2$. The results suggest that the majority (approximately 63 %) of villages are in the low prevalence group (mean prevalence 0.037) with the remainder in the high prevalence group (mean prevalence 0.24). Three possible explanations are:

That these results represent the normal, stochastic behaviour of an infectious disease, like foot-and-mouth disease, in a population. In the many cases, following the introduction of one or a small number of infected individuals, there will be subsequent onwards

transmission of disease to susceptible individuals. In a naive population the typical number of direct onwards cases would be represented by the epidemiological parameter R_0 — the basic reproduction number [Anderson and May, 1992]. If R_0 is close to unity in a naive population or if the effective reproduction number resulting from, for example vaccination, is close to unity, some infection events may result in marked spread of disease within the group of animals but the majority will result in only low numbers of subsequent infections. Hence the two components of the mixture model represent take-off of disease and low transmission with extinction.

An alternative hypothesis is that the mixture arises not, primarily, as a result of stochasticity but that the effective reproduction number differs between groups of villages. The effective reproduction number in a Thrace village will vary according to vaccination coverage, virus strains, environment and many features of husbandry practice. It is thus possible to hypothesise that the mixture distribution of prevalence within villages is a consequence of, for example, high and low vaccination coverage or differing husbandry practices.

It is also possible that the mixtures actually represent a high prevalence distribution of truly exposed villages where infection has taken off and a low apparent prevalence distribution which arises due to poor specificity of the diagnostic test. This explanation, however, is not supported by the parameter estimates for the test specificity which were, as noted previously, robust to diagnostic test prior assumptions.

To identify the process behind the apparent mixture distribution requires accurate, village by village vaccination coverage estimates and in situ evaluations of the CEDITEST diagnostic performance parameters when used in the Turkish laboratories on sera from Turkish (vaccinated and unvaccinated) cattle.

5.5.2 Diagnostic test performance

The diagnostic test was estimated as slightly less sensitive (85%) than the design sensitivity used in the Thrace surveys (90%) but was highly specific (99.85%) and close to the design specificity of 100%. Figure 5.2 shows that the posterior distribution of the sensitivity was largely driven by the prior from Engel et al. [2008] whereas the posterior for specificity was updated by the model/data to be more specific with less uncertainty. Qualitatively the sensitivity and specificity estimates arise because for a given model

the likelihood of observing a given data-set will be dependent on the sensitivity and specificity of the diagnostic test (just as it is dependent on the village and animal-level prevalence). Thus given a prior distribution, the Bayesian approach potentially allows estimation of the test parameters. However there almost no information to inform the sensitivity of the diagnostic test as this parameter will be negatively correlated with prevalence estimates and there are low numbers of test positive animals. Hence the posterior for the diagnostic test sensitivity is almost identical to the prior. The specificity is more strongly identified by the model as a marked departure from perfect specificity will effectively introduce a third mixture distribution into the model (animals will test positive because they are falsely positive or because they are genuinely positive in a village of one of the two animal-level prevalence groups).

5.5.3 General assumptions

The model was constrained to fix diagnostic test sensitivity and specificity as constant across all three surveys and to be fixed within each survey. The earlier analysis of percentage inhibition results (in chapter 4) suggested that this is not necessarily the case. The apparent drift in location of percentage inhibition was corrected by re-calibration of the test (effectively moving the cut-off point) as previously discussed. However, it is possible that other effects that do not move the overall location of the percentage inhibition test results may occur between and within surveys and would cause a between survey variation in test performance. Unfortunately there are insufficient observations to attempt to estimate sensitivity and specificity on a survey by survey basis. Trial attempts at this resulted in non-convergent models. This is unsurprising as the information content of the data regarding test sensitivity will be very low with the low rate of positive test results.

Also as previously discussed the model uses a simplification that the prevalence within a sample is the prevalence within a village. This assumes that the sample is a representative sample from the village cattle population. This is not the case. In order to reduce the risk that cattle having multiple previous foot-and-mouth disease vaccinations will be tested positive, due to vaccines containing small amounts of non-structural proteins, animals for surveillance are selected from those between 4 months and 24 months old. Thus the prevalence estimates for foot-and-mouth disease exposure will be biased representing the prevalence in this age group which may have a different exposure risk and

behaviour to the average over the whole population a village. It is also possible that sampling will not occur perfectly at each site as sampling is locally delegated (Naci Bulut, ŞAP Institute, personal communication 2006) and hence may occasionally be on a convenience basis from single farms rather than randomly across a whole village sample frame.

Follow up CEDITEST diagnostic tests were performed by the ŞAP institute on positive animals with then further follow up using the Bommeli CHEKIT test. I have elected to only analyse the initial test results rather than use the results of the follow up testing. This was because only limited information is available describing the conditional interpretation of these tests [Paton et al., 2006]. Hence inclusion would require joint estimation of these parameters in the Turkish setting and would be unlikely to add to the precision of the epidemiology estimates.

5.5.4 Implications for sero-surveillance in Thrace

As shown in Chapter 3, surveys to demonstrate freedom from disease using tests with significantly imperfect specificity will almost always fail to demonstrate freedom on *initial* interpretation of the primary test results. The simple pass/fail results from the raw data of the 2005 and 2006 surveys in Thrace surveys has no utility for charting the progression of a region from endemic or epidemic status to disease freedom as, by strict interpretation, the region would be considered 'not-free' at each survey. Quantitative analysis, such as the Bayesian mixture model approach, may permit a finer grained examination of the primary NSP ELISA test data, before application of confirmatory tests, and allow inference about the epidemiology of disease and its trends. The key results from such analyses are the estimates of village-level prevalence of exposure ('*v.prev*' in this study) and the estimates of within village prevalence. These describe the village-level of incursion of disease or previously exposed livestock and the within-village characteristics of these incursions.

The analysis of the three surveys from Thrace suggests that in all surveys some villages contained animals that were previously exposed to foot-and-mouth disease virus. It can not be definitively ascertained from the data if this represents exposure of animals within the village or movement into the village of previously exposed livestock. There appeared to be a trend in estimated village-level prevalence of exposure compatible with

reported disease history although higher in magnitude. This may represent undetected disease, may be a result of vaccination related effects on test specificity or be a direct result of diagnostic test performance. It would be important to identify which, by further evaluation of the diagnostic test in the relevant vaccinated cattle, before a region under vaccination control could be demonstrated to be free of FMD solely by NSP test based sero-surveillance without further confirmatory testing.

More generally, the use of quantitative results in charting the progression to disease freedom is likely to require careful refinement of the epidemiological model to reflect the mechanics of incursion, movement, sampling and diagnosis. In the Thrace setting a low level of seropositivity at a village level may be unavoidable due to purchase of previously exposed livestock and the effect of residual traces of NSP material in vaccines. Quantitative models need to explicitly include these processes otherwise villages may be estimated as exposed to FMD when they are not.

5.6 Appendix

5.6.1 Model code

5.6.1.1 Base two-level model

$$
\left.
\begin{aligned}
\text{pos}_i &\sim \text{dbin}(\text{ap}_i, \text{samples}_i) \\
\text{ap}_i &= \text{se} \cdot p_i + (1 - \text{sp}) \cdot (1 - p_i) \\
p_i &= p_i^{\star} \cdot v_{\text{status},i} \\
p_i^{\star} &\sim \text{dbeta}(a_{a,1}, b_{a,1}) \\
v_{\text{status},i} &\sim \text{dbern}(v_{\text{prev,survey}_i})
\end{aligned}
\right\} 1 \leq i \leq N
$$

$$
\begin{aligned}
a_{a,1} &= \mu.1 \cdot \phi.1 \\
b_{a,1} &= \phi.1 \cdot (1 - \mu.1) \\
\mu.1 &\sim \text{dunif}(0.001, 0.999) \\
\phi.1 &\sim \text{dgamma}(\alpha, \beta)
\end{aligned}
$$

$$
\left.
\begin{aligned}
r_{\text{status},s} &\sim \text{dbern}(r_{\text{prob},s}) \\
v_{\text{prev}^{\star},s} &\sim \text{dbeta}(v.\text{prev}.a, v.\text{prev}.b) \\
v_{\text{prev},s} &= v_{\text{prev}^{\star},s} \cdot r_{\text{status},s} \\
r_{\text{prob},s} &\sim \text{dunif}(0.001, 0.999)
\end{aligned}
\right\} 1 \leq s \leq 3
$$

$$
\begin{aligned}
v.\text{prev}.a &= \mu_v \cdot \phi_v \\
v.\text{prev}.b &= \phi_v \cdot (1 - \mu_v) \\
\mu_v &\sim \text{dunif}(0.001, 0.999) \\
\phi_v &\sim \text{dgamma}(\alpha, \beta) \qquad \alpha = 1 \\
\text{se} &\sim \text{dbeta}(\text{se}_a, \text{se}_b) \qquad \beta = 0.01 \\
\text{sp} &\sim \text{dbeta}(\text{sp}_a, \text{sp}_b) \\
p.a.sim.1 &\sim \text{dbeta}(a_{a,1}, b_{a,1}) \\
v.\text{prev}.sim &\sim \text{dbeta}(v.\text{prev}.a, v.\text{prev}.b)
\end{aligned}
$$

5.6.1.2 Two level model with beta mixture of exposed village's prevalence

$$
\left.
\begin{aligned}
\text{pos}_i &\sim \text{dbin}(\text{ap}_i, \text{samples}_i) \\
\text{ap}_i &= \text{se} \cdot p_i + (1 - \text{sp}) \cdot (1 - p_i) \\
p_i &= p_i^{\star} \cdot v_{\text{status},i} \\
p_i^{\star} &\sim \text{dbeta}(a_{a,\text{class}_i}, b_{a,\text{class}_i}) \\
\text{class}_i &= \text{class}_i^{\star} + 1 \\
\text{class}_i^{\star} &\sim \text{dbern}(p.2) \\
v_{\text{status},i} &\sim \text{dbern}(v_{\text{prev,survey}_i})
\end{aligned}
\right\} 1 \leq i \leq N
$$

$$
\begin{aligned}
p.2 &\sim \text{dunif}(0.001, 0.999) \\
a_{a,1} &= \mu.1 \cdot \phi.1 \\
\mu.2 &= (\text{scale}.2 + \mu.1)/(1 + \text{scale}.2) \\
a_{a,2} &= \mu.2 \cdot \phi.2 \\
\text{scale}.2 &\sim \text{dbeta}(1, 1) \\
b_{a,1} &= \phi.1 \cdot (1 - \mu.1) \\
b_{a,2} &= \phi.2 \cdot (1 - \mu.2) \\
\mu.1 &\sim \text{dunif}(0.001, 0.999) \\
\phi.1 &\sim \text{dgamma}(\alpha, \beta) \\
\phi.2 &\sim \text{dgamma}(\alpha, \beta)
\end{aligned}
$$

$$
\left.
\begin{aligned}
r_{\text{status},s} &\sim \text{dbern}(r_{\text{prob},s}) \\
v_{\text{prev}^{\star},s} &\sim \text{dbeta}(v.\text{prev}.a, v.\text{prev}.b) \\
v_{\text{prev},s} &= v_{\text{prev}^{\star},s} \cdot r_{\text{status},s} \\
r_{\text{prob},s} &\sim \text{dunif}(0.001, 0.999)
\end{aligned}
\right\} 1 \leq s \leq 3
$$

$$
\begin{aligned}
v.\text{prev}.a &= \mu_v \cdot \phi_v \\
v.\text{prev}.b &= \phi_v \cdot (1 - \mu_v) \\
\mu_v &\sim \text{dunif}(0.001, 0.999) \\
\phi_v &\sim \text{dgamma}(\alpha, \beta) \\
\text{se} &\sim \text{dbeta}(\text{se}_a, \text{se}_b) \\
\text{sp} &\sim \text{dbeta}(\text{sp}_a, \text{sp}_b) \\
p.a.sim.1 &\sim \text{dbeta}(a_{a,1}, b_{a,1}) \\
p.a.sim.2 &\sim \text{dbeta}(a_{a,2}, b_{a,2}) \\
v.\text{prev}.sim &\sim \text{dbeta}(v.\text{prev}.a, v.\text{prev}.b)
\end{aligned}
\qquad
\begin{aligned}
\alpha &= 1 \\
\beta &= 0.01
\end{aligned}
$$

5.6.2 Source data

Uses data from Excel spreadsheets from Naci Bulut SAP Institute, Ankara, Turkey

TRAKYA SEROSURVEY 2005.xls 2005 Survey

TRAKYA SEROSURVEY final 12.07.2006.xls Spring 2006 Survey

TRAKYA SEROSURVEY 2006 sonbaharx.xls Autumn 2006 Survey

CEDI SONBAHAR 2006 POSITIVE.xls Positive results from Autumn 2006

villages list and animal pop. size.xls Sampling Frame for 2005 Surveillance

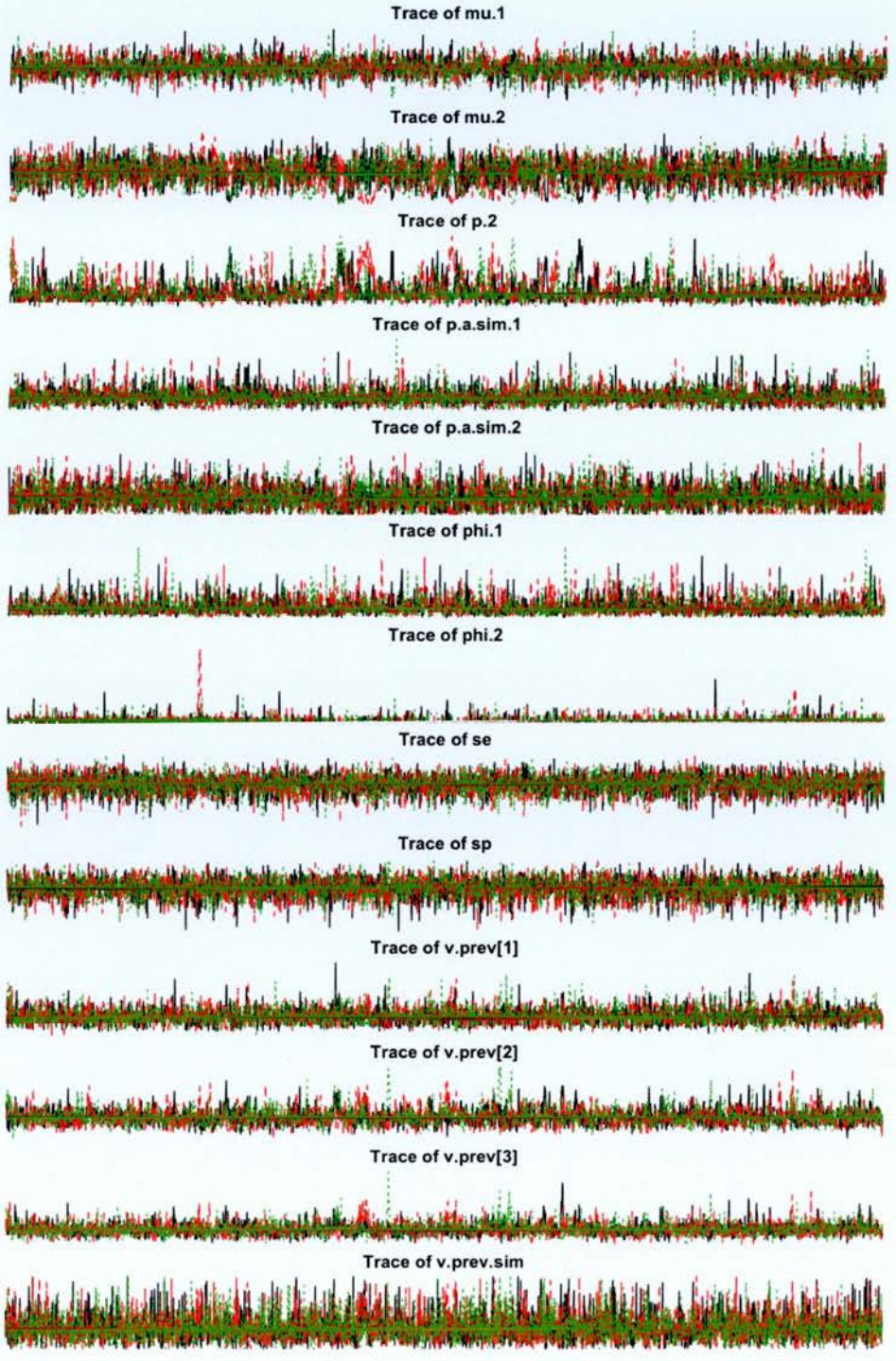
vaccination_report.xls Planned and executed vaccination numbers by district dated March 2006

villages list and animal pop. size.xls List by province, district and village of cattle population census

5.6.3 Trace of model output from JAGS

See following page ...

Figure 5.9: Posterior density trace using zero/double mixture model with re-calibrated data — priors from Engel 2008 for sensitivity and specificity. Three chains are show for each parameters (coloured red, green and black). The samples are thinned displaying 1 sample for every 200 original simulations.



Chapter 6

Clinical versus serological surveillance

6.1 Summary

Current legislation and animal health codes require that after an animal disease epidemic, in a previously free region, several steps are taken to demonstrate that the disease has been controlled and that external trade with the region no longer poses a risk. The two main surveillance activities required are clinical and serological surveillance. Clinical surveillance entails the periodic examination of remaining livestock by stockmen and veterinary surgeons on farm and staff, inspectors and official veterinary surgeons at abattoirs, lairages and markets. Serological surveillance consists of survey-based sampling of animals and subsequent testing with a diagnostic test to identify infected or previously exposed animals. In this chapter I will estimate the diagnostic benefits of the two approaches and estimate the additional value of serological surveillance over clinical surveillance.

Clinical surveillance provides information on the basis that every time stock in a region are examined and signs of infection/disease are not observed the subjective probability that the region is disease-free is greater than before the observation occurred. In this chapter I use this idea formally, applying Bayes theorem, to estimate over the post epidemic period how the probability of disease freedom in a group of livestock increases with time given that stock are inspected from time-to-time and no disease is detected.

I apply this model to post epidemic scenarios using livestock demography data from Devon and Surrey.

The estimation requires a model of observer detection of disease in an affected group of animals. There are three components to this:

- A model of disease epidemiology within the group of animals: I use a stochastic SEIR model with parameters obtained from a literature review and discussion with experts. This captures the stochastic behaviour of foot-and-mouth disease in livestock groups.
- A model of clinical signs of disease: I use a literature review and expert consolation based disease profile which describes the development and resolution of clinical signs throughout the disease course. This is used to approximate the probability with which signs will be detected if an animal is observed.
- A model of stockholder and animal health professional observation: This simulates the periodic and partial observation behaviours that apply to different husbandry systems e.g. beef and dairy cows. A stochastic model was used. This model was informed by estimates of human observation behaviour provided by interviews with veterinary surgeons, stockmen and university clinicians involved in livestock management and disease control activities during the UK 2001 foot-and-mouth disease epidemic.

The performance of serological surveillance has previously been estimated by using relatively simple models of within-group epidemiology and point or distributional models of diagnostic test sensitivity. In my analysis I am particularly interested in the time dependent effects of surveillance so I have explicitly modelled the antibody response of exposed animals against time to give an estimate of diagnostic test sensitivity against time for each animal. Using standard sero-surveillance sampling designs and the same stochastic disease model that I used for clinical modelling allows me to construct a model of serological detection at any time during the post epidemic period.

With the above models I can estimate (for a given *prior* probability of disease in the region at the start of the surveillance period) the daily probability of disease given that no clinical signs are seen. Application of the serological model to the herd or flock simulation on a given date allows me to estimate the further reduction in probability of

disease given that the survey was negative. From these results summary measures can be calculated estimating the time saved by performing serological surveillance after a period of clinical surveillance rather than waiting for a longer period doing just clinical surveillance.

6.2 Introduction and background

After a foot-and-mouth disease epidemic in a previously disease free country international trade regulations [OIE, 2005b] require that evidence is provided to demonstrate that the disease has been successfully controlled and that trade will not represent a significant risk. Conventionally this evidence consists of several components: the methodology of disease control, the epidemiology of the disease, the demography of the affected region, a suitable period of clinical surveillance and a systematically planned and executed sero-surveillance study. This chapter focuses on the last two, surveillance-based, components. Surveillance studies operate by having some, non-zero, probability of detecting disease if it is present in the study population. This is the *system sensitivity* [Martin et al., 2007b]. Completion of the surveillance activity without detection of disease effectively increases confidence that there is no disease present. I use the term *disease* in this study to include all clinical and serological manifestations of disease having potential consequences either clinically or economically. A serologically positive animal may have recovered from clinical *disease*. Detection of such an animal in the post epidemic period would at the very least require its slaughter, may delay declaration of disease freedom and would instigate further diagnostic procedures. Hence for the purposes of this study on post-epidemic surveillance *disease* will include serological as well as clinical signs. Every time an animal, farm or region is examined and disease is not found the probability that the animal, farm or region is disease free increases (assuming the examinations are not completely statistically dependent).

Currently both clinical surveillance — the time to time examination of standing livestock by keepers and veterinary surgeons — and serological surveillance are required as evidence to substantiate freedom from foot-and-mouth disease OIE [2005b]. In this chapter I estimate the surveillance benefits of clinical surveillance in post foot-and-mouth disease outbreak scenarios and then estimate the additional surveillance benefits of serological surveillance.

The post epidemic period has a retrospective start. Towards the tail end of an infectious disease epidemic cases will arise, on average, less and less frequently. The post epidemic period starts with the last recorded case. Notification of a new case effectively 'resets the clock' with a new start of post epidemic period. This process repeats until no new cases are notified. With foot-and-mouth disease, livestock in a previously infected area will normally remain under full or partial movement restrictions until there is strong evidence that there is no longer infectious disease present [Anderson, 2002]. Traditionally, post-epidemic surveillance is a period of high vigilance on the part of farmers, stockmen, government veterinary officers and private veterinary clinicians. There is a high societal pressure within the agricultural community neither to start a new outbreak nor to fail to report existing infection according to a livestock veterinary surgeon active during the UK 2001 epidemic (David Black personal communication, August 2008). Any suspected cases on farms will be reported by stockmen or veterinary surgeons to the local animal health office or disease control centre for subsequent clinical investigation and laboratory diagnostic tests.

6.2.1 Previous studies

McLaws et al. [2007, 2006] estimated the farm level sensitivity of clinical surveillance during the active phase of the UK 2001 foot-and-mouth disease epidemic at approximately 97% by comparison of clinically detected farms with serologically identified farms. However, these studies do not estimate the incremental sensitivity of clinical surveillance with repeated daily examination that will be present in the post epidemic period and do not address the issue of post epidemic serological surveillance. Arnold et al. [2007] uses a stochastic model to simulate foot-and-mouth disease epidemics with clinical identification of infected premises and a culling and vaccination strategy. The study estimates the number of residual carrier animals after an epidemic and considers the efficacy of a non-structural protein based diagnostic test in post epidemic surveillance in vaccinated herds. Carpenter et al. [2004] describe a model for a large dairy herd simulating intra-herd spread of disease and clinical detection. The model is designed to estimate events in a large herd North American setting and uses a very conservative estimate of 10 clinical cases before disease is detected. This and the lack of serological surveillance makes the results of this model unsuitable for estimation of

benefits of serological surveillance in a UK setting. Bates et al. [2003] describe a regional intra and inter herd model of foot-and-mouth disease. This model is used for subsequent evaluation of control strategies including culling and vaccination. As with the other studies described in this section there is no implicit consideration of how the performance of clinical surveillance varies over time and the comparative advantage of serological surveillance at the end of a period.

6.3 Methodology

6.3.1 Overview

The aim of this chapter is to estimate the post epidemic surveillance benefits of clinical surveillance and the additional benefits of serological surveillance. In this study the benefits will be calculated as a daily system sensitivity [Martin et al., 2007b] — the probability that a diseased group or collection of groups of animals would be identified by the survey on a given day. This is then cumulated over the post epidemic period for clinical surveillance to give the probability that a diseased group of animals would be detected on *any* day up to that point. This can then be combined with a prior belief in the probability of disease in the group to give a posterior probability of disease if surveillance has not detected disease.

There are four components to this analysis; a model of epidemiology in infected groups, a model of clinical surveillance, a model of serological surveillance and the integration of these components into the cumulative results. These are addressed in turn in the following sections.

6.3.2 Model of epidemiology within an animal group

Figure 6.1 shows a simplified schematic for the clinical and epidemiological stages of an infectious disease in an individual.

The performance of clinical or serological surveillance will be contingent on the underlying epidemiology of the disease that it aims to detect. As change in surveillance performance over time is a key focus of this analysis an individually based mechanistic model is used to simulate progress of disease through an exposed group of animals. Use

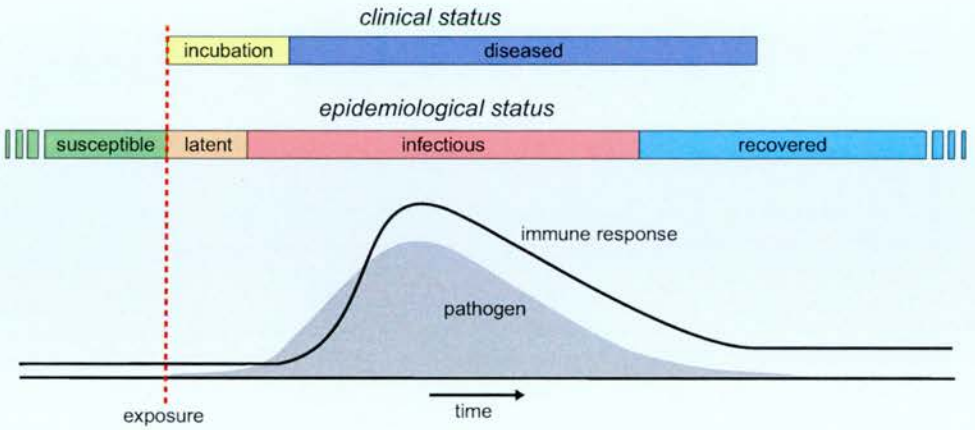


Figure 6.1: Timeline showing pathogen load (e.g. viraemia) and immune response (e.g. antibody levels) of a typical infectious disease in an individual. The coloured bars show the different clinical and epidemiological statuses. (redrawn from [Keeling and Rohani, 2007])

of an individually based model captures both the temporal changes in disease status of animals and the stochastic variability of disease patterns. The model is an SEIR (susceptible, exposed, infectious, recovered) frequency dependent model [Keeling and Rohani, 2007] — a structure which approximates the clinical behaviour of foot-and-mouth disease virus in a dispersed group — where the number of effective contacts between individuals will be approximately independent of the group size i.e. R_0 does not increase as herd or flock size increases [Sobrino, 2004, p 362].

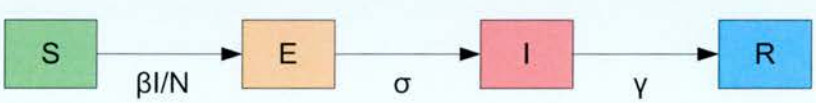


Figure 6.2: Structure of within-group SEIR model showing probabilities that an *individual* changes state. (S=susceptible, E=exposed, I=infectious and R=recovered.) Where β is the product of an individual's average number of contacts per day and the probability of a contact with an infected individual resulting in a successful transmission event, σ and γ are the reciprocals of the latent period and infectious period durations.

At the start of the outbreak the group of animals is assumed to consist of N animals of which E have been exposed to foot-and-mouth disease (by contact, fomite transmission or otherwise) and $S = N - E$ are unexposed i.e. susceptible. The SEIR formulation

uses probabilities that an individual in each class will change state (in the case of the transition from susceptible to exposed this is also dependent on the number of infectious individuals in the whole group). As the transition is represented as a probability rather than a deterministic rate the model is stochastic and the number of individuals in each class with time cannot be determined analytically. A number of methodologies exist to realise this model and generate sets of simulated outbreaks. For efficiency and simplicity I use Gillespie's direct algorithm (Gillespie [1977] described in Keeling and Rohani [2007]). The algorithm uses the following method to exactly simulate the SEIR process:

1. Identify all possible events (e.g. a susceptible individual becomes exposed)
2. Set a time variable to the start time for the simulation
3. Calculate the mean rate at which each possible transition event occurs (e.g. $\beta SI/N$ for $S \rightarrow E$)
4. Calculate the rate at which *any* event occurs (The sum of the individual rates)
5. Draw the time to the next event from the negative exponential distribution using the above rate.
6. Using the individual event rates as weighting choose one of the possible events to perform
7. Do the chosen event and increment the time variable
8. Repeat steps 3 - 7 until the time variable reaches the chosen end of simulation time

The model uses animal group sizes drawn from the agricultural census data for the counties of Devon and Surrey in 2006. Both of these areas have experienced foot-and-mouth disease outbreaks; in 2001 in Devon and in 2007 in Surrey [Anderson, 2002, 2008]. For each simulation of a within-group epidemic the group size is drawn from the empirical distribution of group sizes for the selected county and animal type. The initial number of exposed animals was drawn from a Poisson distribution with a mean of $4.1 + 0.0049h$ where h is the number of individuals in the group. This was based on the initial size estimates of Arnold et al. [2007] in cattle herds in the UK 2001 epidemic.

For my study draws from this distribution that resulted in a zero sized outbreak were discarded. Animal types were divided into sheep, dairy cattle and non-dairy cattle using the census classification data. A farm with dairy animals and non-dairy animals was considered to have two independent groups. The dairy, non-dairy and sheep classes were selected to mirror likely differences in disease epidemiology and clinical observation behaviour between the classes. The distribution of group sizes in the two study counties are shown in the results section. Each within-group outbreak was simulated for 120 days as preliminary simulation studies identified that this period would capture the important differences between sampling methods, scenarios and animal groups. During each simulation the dates of each new transition from susceptible class to exposed class were captured for input into the subsequent stages of the analysis. Illustrative results from the SEIR model are shown in the results section 6.4.2.5.

The SEIR transmission model requires parameters to describe the following: the probability that two individuals contact and that if an individual is infectious a transmission occurs in a unit time (β); the mean duration of the latent period ($1/\sigma$) and the mean infectious period ($1/\gamma$). These parameters were estimated by reviewing the available literature describing the epidemiology of foot-and-mouth disease (see section 6.4.2) and discussion with experts. As there was a large uncertainty associated with these parameters deterministic sensitivity analysis was carried out using different values of β, σ and γ as detailed in the results section.

Limited data were available to validate the SEIR model of within-group infection. One source was the records of post-epidemic serology following the UK 2001 foot-and-mouth disease epidemic. Data were recorded in a report to the OIE for official declaration of disease freedom [DEFRA, 2002]. The records contain details of samples taken from sero-positive farms and the number of animals testing positive. These are summarised in the results section and compared with the outputs of the SEIR model.

6.3.3 Model of clinical surveillance

Clinical surveillance may consist of the visual and physical examination of the livestock in a particular group of animals. This occurs either during their day to day husbandry or additionally as disease control measures. For this study only the observations of livestock keepers were considered. This was both for simulation simplicity and also

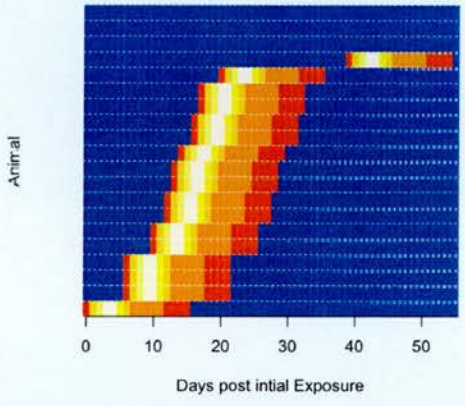
as it was considered that livestock keepers were significantly more likely to identify disease in their stock than external veterinary surgeons or animal health officers as it was suggested by David Black (personal communication August 2008) that they were more familiar with the normal behaviour of their stock and had more opportunity to observe them. A simple model of livestock owner clinical examination was proposed and discussed with an expert panel comprising a foot-and-mouth disease specialist, a production animal clinician, a farm manager, a shepherd and two animal husbandry lecturers.

For each animal type (broadly, their husbandry system):

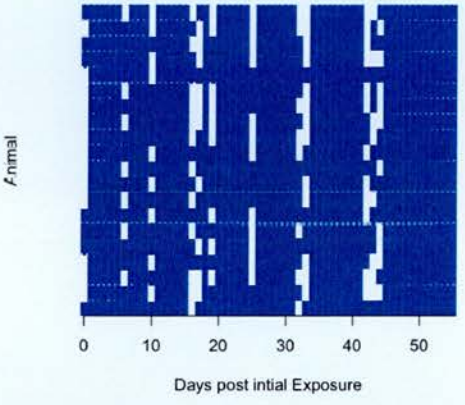
- On a particular day the group of animals may be observed, en masse, with a probability P_{group}.
- If the group was observed each individual would be observed with a probability P_{indiv}.
- If an individual was observed there was a maximum probability of diagnosing foot-and-mouth disease if present P_{max}.
- The relative probability of diagnosis of a foot-and-mouth diseased individual was not constant but followed a clinical disease profile with perhaps subtle signs of disease after an incubation period, then increased signs to a maximum with then a gradual decrease and resolution of signs with recovery and healing.

The parameters for this model were estimated by discussion with the expert panel in the case of the first three parameters and by literature review and discussion with a foot-and-mouth disease epidemiologist in the case of the clinical disease profile parameters.

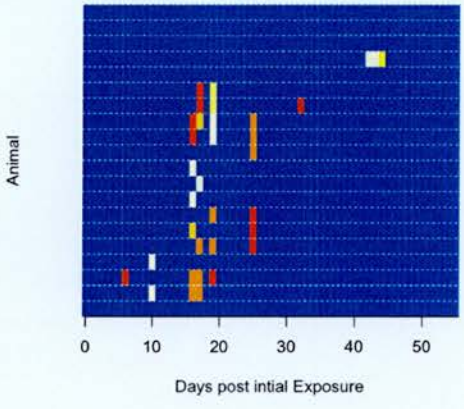
The process of clinical observation was simulated by assuming that each exposed animal in a group progresses through a deterministic clinical profile (see figure 6.3(a)). On each day of the epidemic the probability that the group and an individual within the group was observed was used to create a stochastic (varied on each simulation instance) observation window (see figure 6.3(b)). This observation window was then applied to the clinical profile for the group to give a day by day and animal by animal output of observed clinical signs (see figure 6.3(c)).



(a) Clinical profiles (graded from red-orange-yellow-white in increasing ease of detection)



(b) Stochastic observation window (white = animal observed on that day)



(c) Observed clinical profiles

Figure 6.3: Clinical observation model showing progression of 20 animals through a 60 day outbreak to illustrate use of clinical profiles and observation window. In figure 6.3(a) the colours represent the severity of clinical signs with blue being no signs, orange being low grade signs and white representing severe signs. Figure 6.3(c) represents the outbreak in figure 6.3(a) as observed through the white gaps in the observation windows in figure 6.3(b).

For each animal type (husbandry system) the resulting observed clinical profiles are scaled. For discussion and literature review purposes the clinical presentation of foot-and-mouth disease was estimated on an ordinal scale from 0 (not detectable) to 4 (maximum probability of detection in that animal type/ husbandry system). This scale was assumed to represent a continuous scale. It was standardised so that for each animal type the highest score was P_{max} (the maximum probability of detection in that animal type).

The output of the clinical observation model, applied to simulated epidemic can be represented as a matrix C where $C_{a,t}$ represents the probability that clinical signs will be seen on day t in animal a. An outbreak in a group will be considered detected if clinical signs are detected in *any* animal in the group. Thus the probability of detection of disease in a group of n animals *on* a particular day t can be estimated as:

$$P_t = 1 - \prod_{a=1}^{n} (1 - C_{t,a})$$

Clinical post-epidemic surveillance will begin, by definition, on the first day after an epidemic or outbreak is putatively ended so we are generally interested in the cumulative effect of continued clinical surveillance — i.e. the probability that a diseased group of animals will be detected *by* a given day. This may be estimated at day t by:

$$P_t^* = 1 - \prod_{i=1}^{t} (1 - P_i)$$

P_t may increase or decrease from day to day and in infrequently observed groups of animals may often be zero. P_t^* will always increase or remain constant as time passes.

6.3.4 Model of serological surveillance

Serological surveillance normally consists of the sampling of blood from all or a randomly selected subset of the group of animals for biological testing to identify antibodies to foot-and-mouth disease and hence infer exposure to the virus [Baron, 1996]. In this study I assume that either a random sample of the animal group is selected for serological testing or the entire group is tested. Random sampling is the approach outlined in the DEFRA, EU and OIE guidelines and echoes the procedures carried out in the foot-and-mouth disease epidemic in many regions of the UK in 2001 [DEFRA, 2002]. The mechanism of sample size calculations for such surveys is discussed elsewhere in this thesis and in depth by Cameron [1997]. A standard approach is to calculate a

sample size to give a 95% confidence or probability of the serological test identifying at least one diseased, sampled individual given an assumed within-group prevalence of disease (if it is present) and assumed serological test performance characteristics. In this simulation of serological surveillance I assume that the cELISA diagnostic test [Paiba et al., 2004] (as used in the UK in 2001) is employed. This is a solid phase competitive ELISA test with a theoretical 100% sensitivity in cattle and sheep.

In the simulation model the sample size for the random sample required to meet the above 95% confidence criteria was generated for each simulated epidemic using a look up table of herd/flock size and sample size. The look-up table was created using the software program 'FreeCalc' [Cameron and Baldock, 1998a] with the above diagnostic test parameters and an expected herd prevalence of foot-and-mouth disease (if infected) of 5% as used for serological survey design in 2001 [Thrusfield et al., 2005a].

In this study the diagnostic benefits of clinical surveillance are assessed from the first day of the post epidemic period through to 120 days. It is possible that an animal may be tested with the cELISA test that has not mounted a full serological response and hence the diagnostic test will not operate at maximum sensitivity. Animals exposed to foot-and-mouth disease develop antibodies after a delay and effectively have a serological profile, with time varying probability of detection. The effect of this serological profile on the test performance was estimated using data from the cELISA evaluation in Paiba et al. [2004]. In Paiba et al. [2004] two groups of four cows and two groups of four sheep were inoculated with foot-and-mouth disease. One group in each species was inoculated directly, the other by contact with foot-and-mouth disease infected donor animals. Blood samples were drawn from the exposed animals and the percentage inhibition results of a cELISA assay were recorded. For the evaluation of the variation of diagnostic test sensitivity with time after exposure I have used the data from the group that were indirectly exposed via a donor animal as this is more likely to represent the response of exposed animals in a field setting. The original data from this study were not available so percentage inhibition values for each of the animals were measured from the published figures using digitisation software [Arizona-Software, 2008]. Blood samples were recorded as taken at 0, 2, 4, 6, 8, 10 and 12 days in the sheep and 0, 1, 2, 3, 4, 5, 7, 9, 11, 15 and 30 days in cattle. In order to impute the test behaviour in the intervening days and to estimate the variability of the cELISA response an empirical model was fitted to the recorded data. A four parameter logistic model [Crawley, 2007, page 203] was used to capture the asymptotic and sigmoid features of the antibody

response with time. The four parameter logistic model was chosen as it allows fit to lower and upper asymptotes with flexibility of location and scale of the sigmoid component of the curve. The variance of percentage inhibition results was obviously heteroscedastic with time so process error was modelled as an exponential polynomial function. The percentage inhibition response at time t (PI_t) was modelled:

$$PI_t \sim N(\mu_t, \sigma_t)$$

$$\mu_t = A + \frac{B - A}{1 + e^{C \cdot (D + t)}}$$

$$\sigma_t = 1/\sqrt{\tau_t}$$

$$\tau_t = exp(E + F \cdot t + G \cdot t^2)$$

The model was parameterised using Bayesian MCMC with JAGS software [Plummer, 2003] called using the R statistical environment [R Development Core Team, 2007] using the Runjags package [Denwood, 2008]. The model formulation is shown in appendix 6.6. Vague gamma priors were used for the parameters A..G. The model (see section 6.3.4) was run for a burn-in period of 100,000 simulations and sampled for 100,000 iterations. At each day post exposure over the range of the evaluation data the percentage inhibition values were simulated (within the MCMC code) and compared to the manufacturer's cut-off value of 60% (the value that would be used to classify results in an outbreak). This permitted an estimation of the diagnostic test sensitivity at each day post exposure. Convergence of the MCMC process was assessed by running three simulation chains and examining the posterior traces of these chains and calculating the Gelman convergence diagnostic [Brooks and Gelman, 1998]. As this model was simply to smooth the underlying cELISA data and capture its uncertainty no additional model diagnostics or alternative models were assessed.

The model estimates the sensitivity of the cELISA test applied to indirectly infected cattle and sheep for each day after their initial exposure. These results are used, in a similar manner to the clinical sign profiles in the previous section, to give a serological detection probability profile (sensitivity) to each animal in the group as it becomes exposed. The exposure times are exactly the same times as simulated by the stochastic SEIR model. This whole group profile is then randomly sampled on each day (using

the sample size determined by FreeCalc) or selected in full depending on the sampling strategy being simulated to give a simulated set of blood results for each, post-epidemic, day. The probability that one or more of these animals will test positive is given by:

$$P(T^+) = 1 - \prod_{i=1}^{s} (1 - Se_i)$$

where $P(T^+)$ is the probability of at least one animal testing positive and Se_i is the estimated sensitivity of the test in the i^{th} animal in the sample of s animals.

Serology is unlikely to be repeated so a cumulative sensitivity of testing the group is not calculated.

For both the serological surveillance and the clinical surveillance only the survey *sensitivity* is estimated. With both modes of surveillance there may be false positive results as neither surveillance mode will have perfect *specificity*. This is not considered in this analysis as I am concerned with the relative ability of the different survey modes to detect potentially diseased animal groups. False positive results will be screened out after clinical or serological surveys by further clinical, serological and virological investigations involving examination by experts and additional diagnostic tests.

6.3.5 Estimate the posterior probability of disease

The above methods allow the estimation of the cumulative, day-by-day sensitivity of clinical surveillance $P^*(T^+)_t$ and the sensitivity of serological surveillance $P(T^+)_t$ at day t. Given a prior belief of the probability that a group of animals is infected the clinical surveillance sensitivity results can be used to estimate the probability that the group is infected day by day during the post epidemic period given that no signs are clinically detected. Additionally the further reduction in probability of disease given a negative serological survey result can also be estimated on a day at the end of a period of clinical surveillance. In qualitative terms: *for every day that disease is not detected we become more sure that it is not present and if we additionally perform serology we are further reassured that the group of animals is not infected.*

$$Prob(D^+|T^-) = \frac{Prob(T^-|D^+) \times Prob(D^+)}{Prob(T^-)} \tag{6.1}$$

So an initial prior belief that a group of animals is infected of $Prob(D^+)$ is updated using the probability of clinical detection up to day t to give a posterior belief that a group of animals is infected (given no clinical disease was observed) of $Prob(D^+|T^-_{clin})$.

If on a particular day clinical surveillance is done the posterior estimate of the probability of disease to that point ($Prob(D^+|T^-_{clin})$) can be used as the prior estimate for a similar equation to 6.1. This gives the probability of disease in the group given that no clinical signs of disease have been observed up to that day and that serological surveillance on that day was negative.

6.3.6 Summary metrics

For each simulation set summary measures are calculated. Generally the posterior probability of disease will approach an asymptote whereby further clinical surveillance with no disease detected does not further decrease the posterior probability of disease. This point is estimated numerically by identifying the first date at which the posterior probability of disease is within a specified tolerance (1% of the prior probability of disease). This probability of disease or confidence of disease freedom could also be achieved by doing serological surveillance at the end of a period of clinical surveillance. The date of this can be identified and hence a number of 'days saved' estimated to achieve the same confidence of disease freedom. An alternative measure is to take the asymptote point of the clinical surveillance result and estimate the benefits of additional serological surveillance at this date giving 'probability of disease reduction'. See figure 6.4 for an illustration of these measures. The robustness of the summary measures to the tolerance method was assessed using deterministic sensitivity analysis setting the asymptote detection tolerance to 2% and 0.5%.

6.3.7 Prior probability of disease

The previous section requires a prior probability of disease for the calculation (using Bayes theorem) of the posterior probability of disease each day given negative survey results. This probability will be informed on an individual basis by factors including epidemic history and control strategy. The base figure used for this analysis is 0.0002. This is the equivalent of one farm in Devon being diseased (there being approximately 5000 cattle/sheep groups in Devon). DEFRA's report to the OIE claiming disease

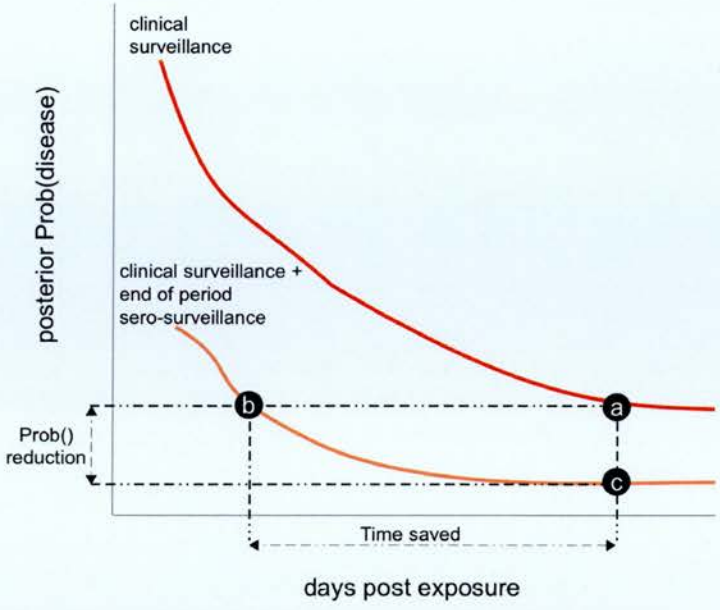


Figure 6.4: Illustration of measures derived from the model outputs. The red (upper) line represents the decreasing probability of disease as time passes with no clinical detection. The orange (lower) line represents the additional benefits of carrying out serological surveillance at the end of a given period of clinical surveillance. Points (a), (b) and (c) are estimated numerically to calculate the two metrics - 'time (or days) saved' and 'probability reduction'.

freedom after the UK 2001 foot-and-mouth disease epidemic [DEFRA, 2002] identified that 0.03 % of farms that were serologically tested were test positive. This estimate may represent an upper limit of diseased farms after that epidemic as some of these farms will have been old infections missed during the main epidemic control period. The values 0.001 and 0.0001 were used for sensitivity analysis.

6.4 Results

6.4.1 Demography of study areas

The 2006 agricultural census records 8,796 holdings in Devon and 2,265 in Surrey. These are classified by presence of any cattle or sheep in tables 6.1 and 6.2. The agricultural census classifies animals on a holding into categories, distribution of group sizes in the two counties within these categories is shown in figures 6.5. For the simulation models of outbreaks and surveillance cattle were considered to be either in dairy or

non-dairy (hereafter referred to as 'beef'). This grouping was chosen as it reflected the broad difference in stockman observation behaviours between the twice daily close observation of milking dairy cattle and the less frequent, longer range observation of non-milking cattle. Dry cows, heifers, bulls, calves and other non-milking stock on dairy units were placed in the 'beef' category.

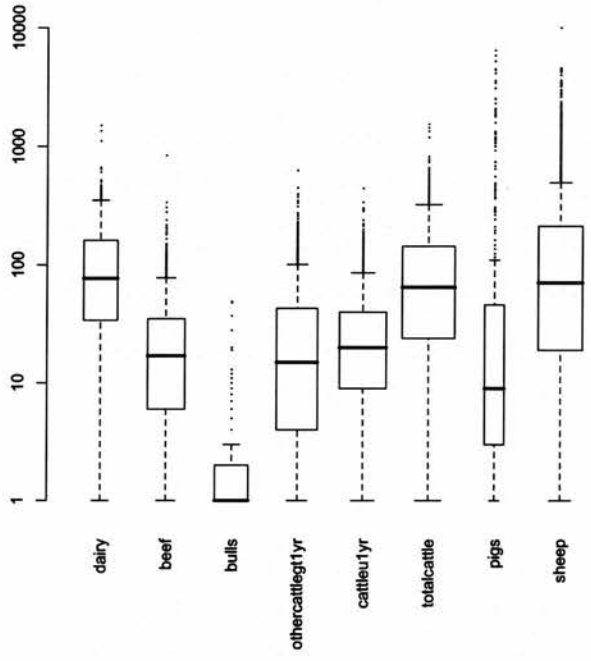
	No Cattle	Cattle
No sheep		1839
Sheep	1031	919

Table 6.1: Number of farms in Devon with either any cattle, any sheep or both

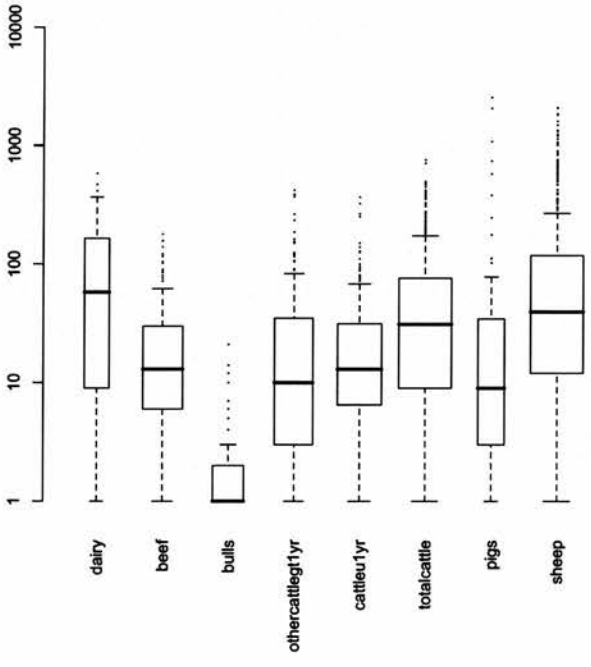
	No Cattle	Cattle
No sheep		295
Sheep	297	129

Table 6.2: Number of farms in Surrey with either any cattle, any sheep or both

There is a marked difference in the total numbers of cattle and sheep holdings in the two counties. Additionally there are relatively fewer dairy herds in the 1–200 animal herd size range in Devon compared to Surrey. The distributions of classified group sizes in the two counties are shown in figure 6.6.

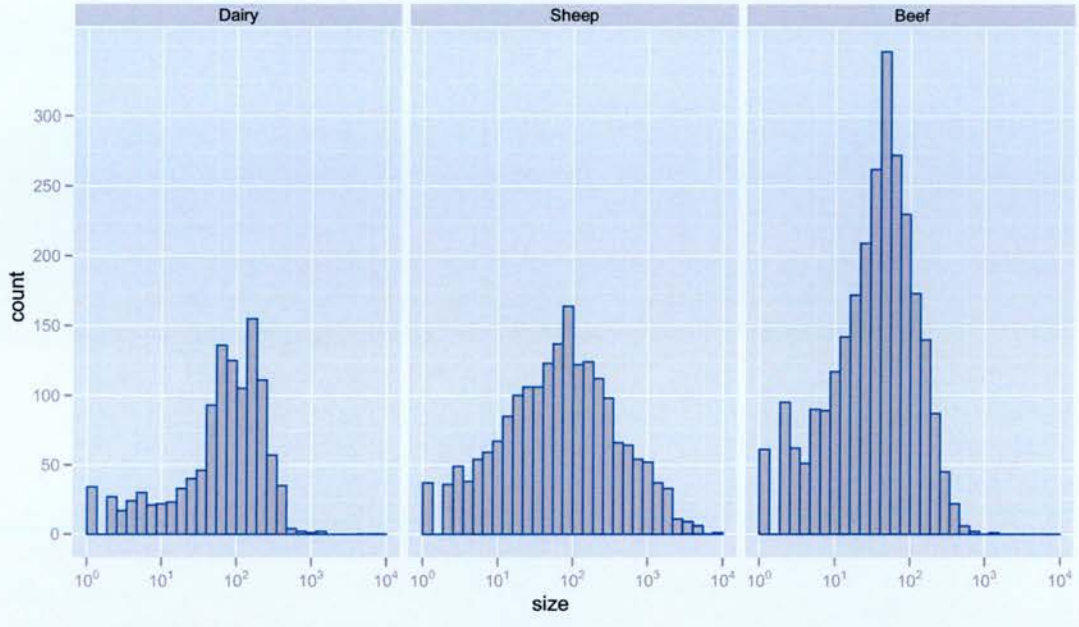


(a) Devon

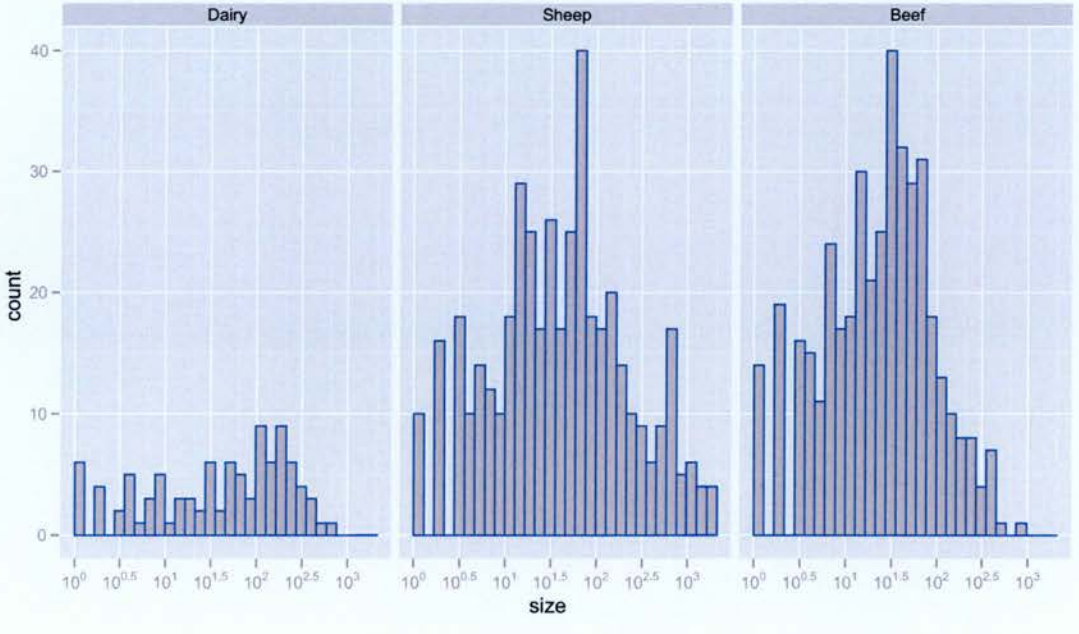


(b) Surrey

Figure 6.5: Box plot showing the distribution of herd and flock sizes as classified by demography records in the 2006 Agricultural census. Herd and flock sizes are shown on the y-axis with group classification on the x-axis



(a) Devon



(b) Surrey

Figure 6.6: Distribution of dairy, beef (non-dairy cattle) and sheep groups in Devon and Surrey scenario areas.

6.4.2 Parameters for SEIR Epidemic Model

6.4.2.1 Transmission parameters

The majority of studies regarding foot-and-mouth disease epidemiology consider the farm or animal holding to be the unit that may become infected ([Ferguson et al., 2001; Keeling et al., 2001; Keeling et al., 2003; Keeling and Rohani, 2007; Tildesley et al., 2006] and others). This is appropriate from the point of view of farm based decision making such as vaccination strategies, movement controls and culling policies. However my model of within farm surveillance requires transmission parameters describing animal to animal transmission characteristics. These are generally estimated experimentally [Velthuis et al., 2007] for reasons of control and reproducibility although they may be estimated from surveillance data obtained by clinical examination or diagnostic testing of farm animal groups. Estimates from the literature review are summarised in table 6.3 with notes in section 6.7. There is great uncertainty in some published studies regarding the transmission parameters of foot-and-mouth disease at an animal-level in within-farm outbreaks. For example in Orsel et al. [2007a]'s study of transmission between dairy cows the estimate for R_0 in unvaccinated animals had a 95% confidence interval of 1.2 to infinity.

6.4.2.2 Latent Period

Carpenter et al. [2004] provide estimates for the latent period and infectious period in dairy cattle based on expert opinion and literature review. Bates et al. [2003] provide an estimate for the latent period in sheep (see table 6.3).

6.4.2.3 Infectious Period

Alexandersen et al. [2003]'s analysis of virus concentration in excretions from infected animals demonstrated virus excretion at least 10 days after exposure of cattle to infected pigs. However they report a sharp decline in excreted virus load at day 4–5 of clinical disease. Orsel et al. [2007b] estimated the infectious period (using an SIR model without a latent period) as 21.1 days with a 95% confidence interval of 10.6–42.2. Parker [1971] estimated shedding times in cattle to be from 1 to 12 days depending on strain (using

Study	Description	R_0 or β	Latent Period(days)	Infectious Period(days)
Orsel et al. [2005]	Calves	R_0 2.52 (1.13–52.1)	—	—
Orsel et al. [2007a]	Dairy cows	R_0 ∞	—	—
Orsel et al. [2007b]	Lambs	β 0.105 (0.044–0.253)	—	21.1 (10.6–42.1)
Orsel et al. [2007b]	Lambs	R_0 2.22 (0.46–4.33)	—	21.1 (10.6–42.1)
Carpenter et al. [2004]	Dairy cows		3.7 (2.13–5.27)[1]	21.0[2] (3.9–53.4)
Bates et al. [2003]	Sheep		6.62 (3.97–10.98)[1]	
Parker [1971]	Dairy Cattle			range 1–12
Streftaris and Gibson [2004]	Sheep	R_0 0.5763 β 0.024 (0.005–0.048)	1.599 (0.943–2.695)	—
Streftaris and Gibson [2004]	Sheep	R_0 0.4803 β 0.020 (0.004–0.036)	1.937 (0.836–3.403)	—
Bankowski et al. [2008]	Cattle[4]		4.46 (1.68–9.68)	2.22 (1.76-2.99)
Bankowski et al. [2008]	Cattle[5]			2.86 (2.42–3.62)
Mardones et al. [2008]	Cattle		1.8	5.9
Mardones et al. [2008]	Small Ruminants		2.0	2.8
Nsubuga et al. [2008]	Sheep	β 1.07	0.98	2.87

Table 6.3: Transmission parameters and latent / infectious periods from literature review used to parameterise SEIR within-group model (all periods measured in days and intervals are 95% confidence intervals from original studies unless otherwise stated). [1]Derived from distributions stated in publication. [2]Sum of sub-clinical and clinical infectious periods in the study. [3] assuming 8 susceptible animals and generation time of 3 days. [4]Estimated assuming infectious period and exposure have to overlap in experiments. [5]Estimated assuming infectious period has to cover all of exposed period in experiments.

Animal Type	R_0	latent period $(1/\sigma)$	infectious period $(1/\gamma)$
Sheep	0.5	2	3
Beef	2.0	2.5	5
Dairy	2.0	2.5	5

Table 6.4: Estimates used to parameterise SEIR model.

a mouse model to assess infectivity). Alexandersen et al. [2002b, 2003] consider sheep to have a highly infectious period of 7–8 days followed by a period of 1–3 days when trace amounts of viral RNA are excreted and subsequently a carrier state in 50% of individuals. It is also noted that although the majority of animals are no longer secreting by 10–14 days post infection environmental virus survival will effectively prolong the infectious period. This environmental survival can be up to 6 months in slurry although there is a poor evidence base for this [Alexandersen et al., 2003].

6.4.2.4 Selection of parameter estimates for SEIR model

The previous estimates from the literature suggest a high level of uncertainty regarding the values of transmission parameters for foot-and-mouth disease within-groups of animals. This will, in part, reflect the stochastic and husbandry system dependent nature of transmission and clinical disease. The parameter ranges used for simulation were informed by values found in the literature and discussion with an epidemiologist experienced with foot-and-mouth disease. The base values are recorded in table 6.4. Further scenarios were then developed using a range of basic reproduction numbers for the outbreaks and also lengthening the duration of latent and infectious periods. These scenarios are shown in table 6.5. They use values of R_0 of 0.5, 1, 2 and 8 for all groups and include extended scenarios with extended latent/infectious periods for all groups.

6.4.2.5 Validation of SEIR model

The SEIR model was internally validated by comparison of its outputs with those of a deterministic SEIR model using the same parameters. The deterministic model was estimated with a numerical method [Woodrow Setzer, 2008] using the *lsoda* algorithm [Hindmarsh, 1983]. The results of two simulation sets are shown in figures 6.7 and 6.8. There appears to be good agreement between the results of the stochastic SEIR model

Scenario	Animal type	R_0	$1/\sigma$	$1/\gamma$
1	sheep	0.5	2	3
2	sheep	1.0	2	3
3	sheep	2.0	2	3
4	sheep	8.0	2	3
5	sheep	0.5	4	6
6	sheep	1.0	4	6
7	sheep	2.0	4	6
8	sheep	8.0	4	6
1	cattle	0.5	2.5	5
2	cattle	1.0	2.5	5
3	cattle	2.0	2.5	5
4	cattle	8.0	2.5	5
5	cattle	0.5	5	10
6	cattle	1.0	5	10
7	cattle	2.0	5	10
8	cattle	8.0	5	10

Table 6.5: The transmission, latent period and infectious period parameters used for the 8 epidemiological scenarios for each animal type in the SEIR model.

and the deterministic SEIR model suggesting that, for the purposes of this analysis, the stochastic model is correctly simulating the intended transmission process.

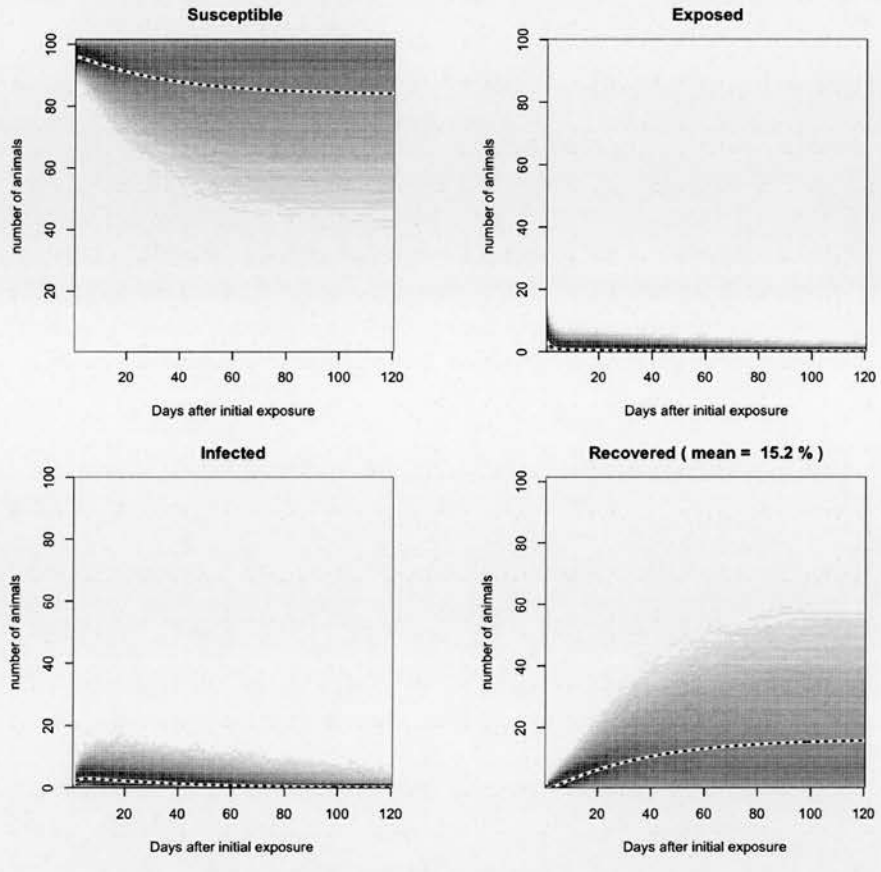


Figure 6.7: An example result set from the SEIR simulation model in a sheep flock of 100 animals. The plots show the results of 50000 simulations using initial exposed numbers as described in section 6.3.2. R_0 is 0.8, the latent period is 4 days and the infectious period is 8 days. The plots show the density of the different outbreak trajectories (trajectories shown white–light grey–dark grey as more frequent). The black dashed line on each plot shows the result of a deterministic outbreak model with the same parameters as the stochastic model.

Ideally, external validation would compare the stochastic SEIR simulation output with detailed records of comparable on-farm epidemics where animals were sampled and tested for foot-and-mouth disease with the husbandry system unaltered. Unfortunately the highly infectious nature, economic importance and regulatory issues surrounding foot-and-mouth disease make such experiments impractical. Whenever foot-and-mouth disease is detected in a herd or flock controls will be applied and normally the animal

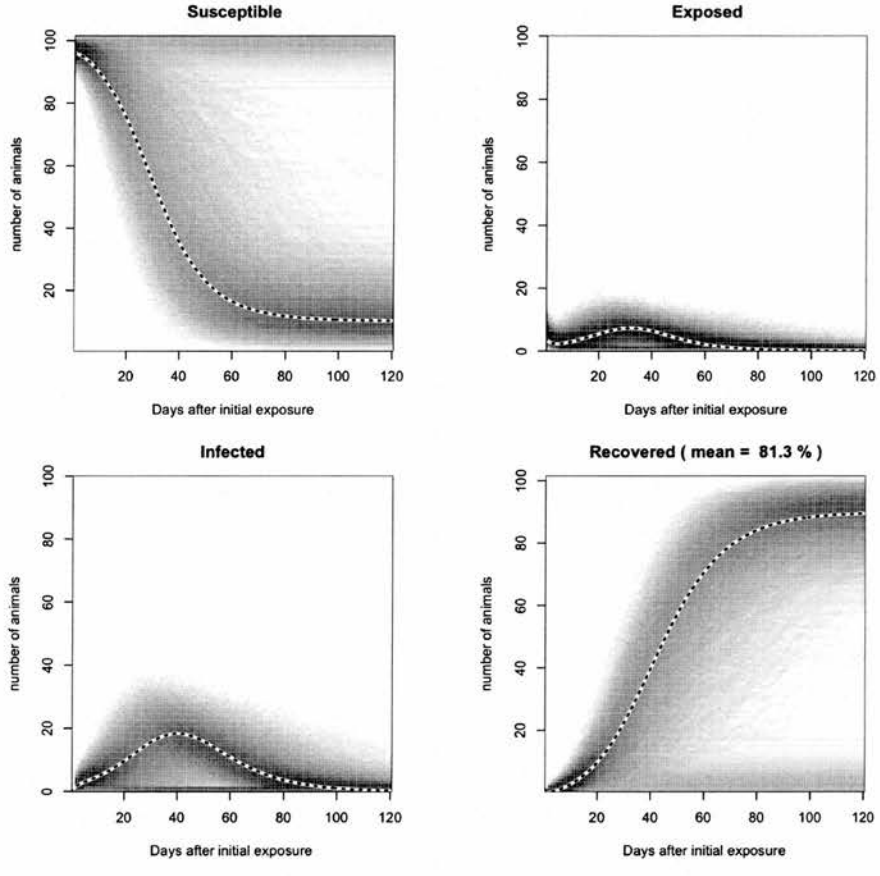


Figure 6.8: An example result set from the SEIR simulation model in a cattle herd of 100 animals. The plots show the results of 50000 simulations using initial exposed numbers as described in section 6.3.2. R_0 is 2.4, the latent period is 4 days and the infectious period is 8 days. The plots show the density of the different epidemic trajectories (trajectories shown white–light grey–dark grey as more frequent). The black dashed line on each plot shows the result of a deterministic epidemic model with the same parameters as the stochastic model.

group will be culled as soon as possible. The DEFRA report to the OIE for demonstration of disease freedom lists records of 45 sero-positive sheep premises identified in the post epidemic surveillance [DEFRA, 2002]. These farms escaped clinical detection through signs not being observed due to poor observation, chance or a less obvious clinical manifestation of disease. These results provide an estimate of epidemic final size without external controls. The estimate may be biased as farms where disease is harder to detect are more likely to contribute to the sample e.g. smaller farms with less experienced stockmen. The point estimates of on farm prevalence are shown in figure 6.9. The mean within farm prevalence was 10.9%. Of the 45 sampled farms 22 had a single serologically positive animal. These results are loosely compatible with the estimated final size prevalence of 8.36% in sheep flocks using the base epidemic parameters as shown in table 6.6

Several studies document prevalence of foot-and-mouth disease within herds of vaccinated cattle [Kitching et al., 2005; Woolhouse et al., 1996] and in endemic settings [Bronsvoort et al., 2006]. Woolhouse et al. [1996] estimated R_0 to be between 2.1 and 72.8 in 5 different herds with a median of 21.1. These were large(1732–2834 head) cattle herds with outbreaks 71–114 days after vaccination. Bronsvoort et al. [2006] reports within-group prevalence estimates for cattle of 34.2%, 83.6% and 91.1% for serotypes O, A and SAT2 respectively. For sheep and goats estimates were 61.1% and 5.6% for serotypes SAT2 and O respectively. These estimates are based on serology of Cameroonian livestock in the absence of vaccination or culling based control. These estimates are of limited value in validating the SEIR model but do illustrate the high, within herd prevalences that can be seen with foot-and-mouth disease in the absence of control.

For illustration of the different epidemiology scenarios used for sensitivity analysis the results from 10000 simulations of an epidemic in a group of 100 sheep and 100 cows are shown in figures 6.10 and 6.11. The results show the wide variation in both the timings of exposure and the final numbers of exposed animals estimated by the stochastic simulation. Table 6.6 shows the estimated final prevalences under eight different epidemiological scenarios.

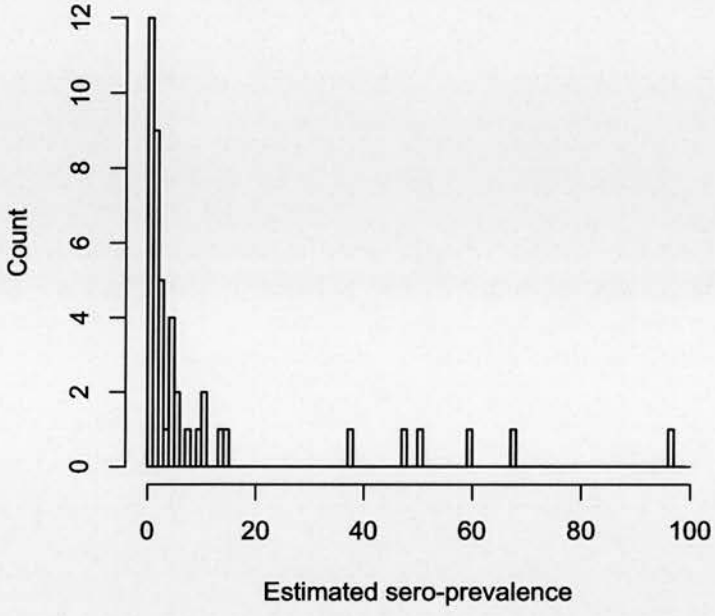


Figure 6.9: Point estimates of within farm sero-prevalence obtained from the 45 sero positive farms recorded in annexe 5 of [DEFRA, 2002]

Scenario	Animal type	Final prevalence (%)
1	sheep	8.4
	cattle	8.5
2	sheep	20.5
	cattle	20.6
3	sheep	69.9
	cattle	69.6
4	sheep	98.3
	cattle	98.2
5	sheep	8.3
	cattle	8.3
6	sheep	20.4
	cattle	19.3
7	sheep	68.9
	cattle	65.7
8	sheep	98.2
	cattle	98.2

Table 6.6: Mean final prevalence at 120 days after initial infection in six transmission parameter scenarios (deterministic sensitivity analysis) for sheep and cows. Result of 50,000 simulated outbreaks.

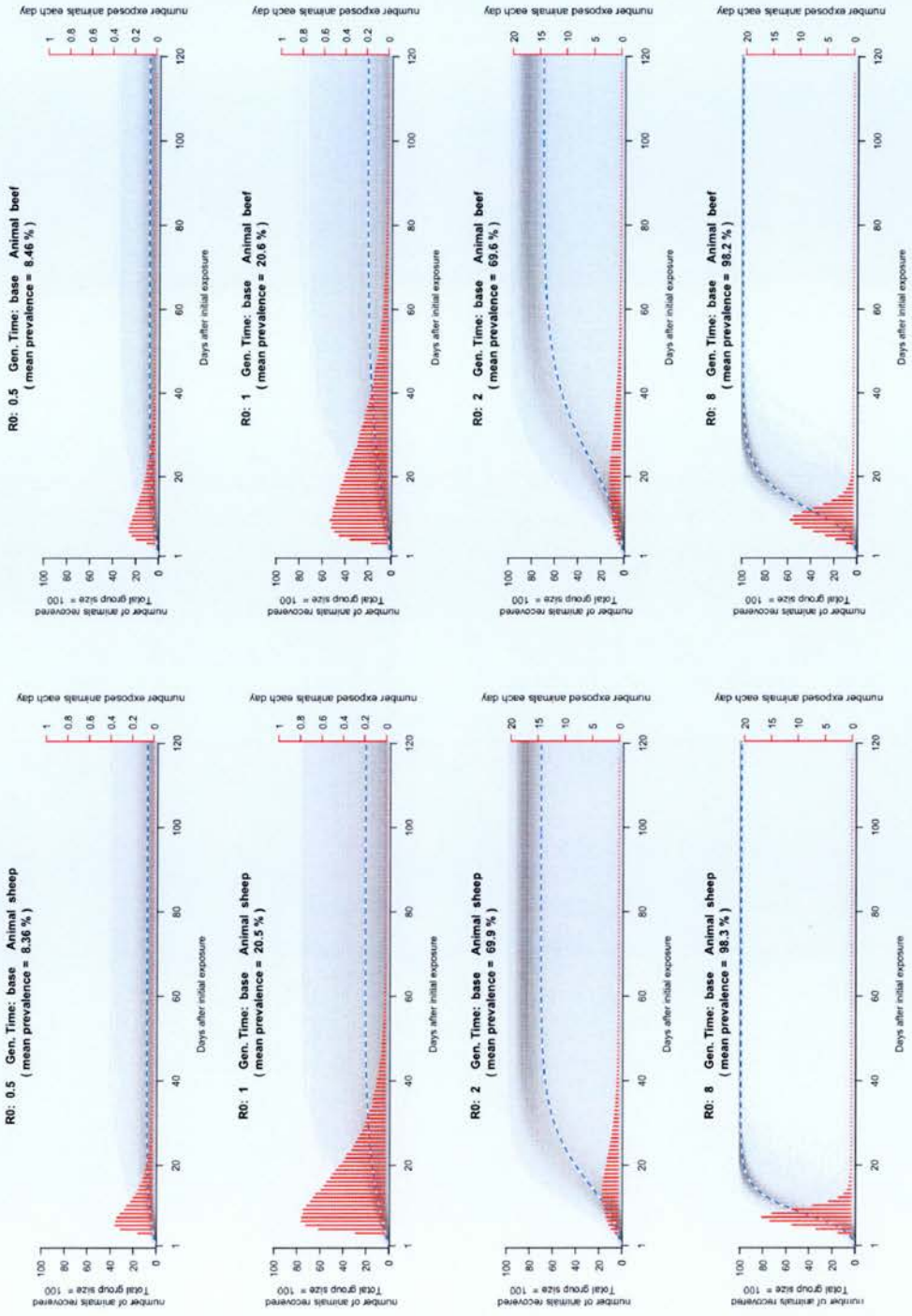


Figure 6.10: Output from the SEIR model for scenarios 1–8 with base incubation and infectious periods showing distribution number of recovered animals against time (grey density plot), mean output from model of recovered animals (blue/white dashed line) and daily new number of exposed animals (red bars). The plots are the result of 10,000 simulations in a group of 100 animals.

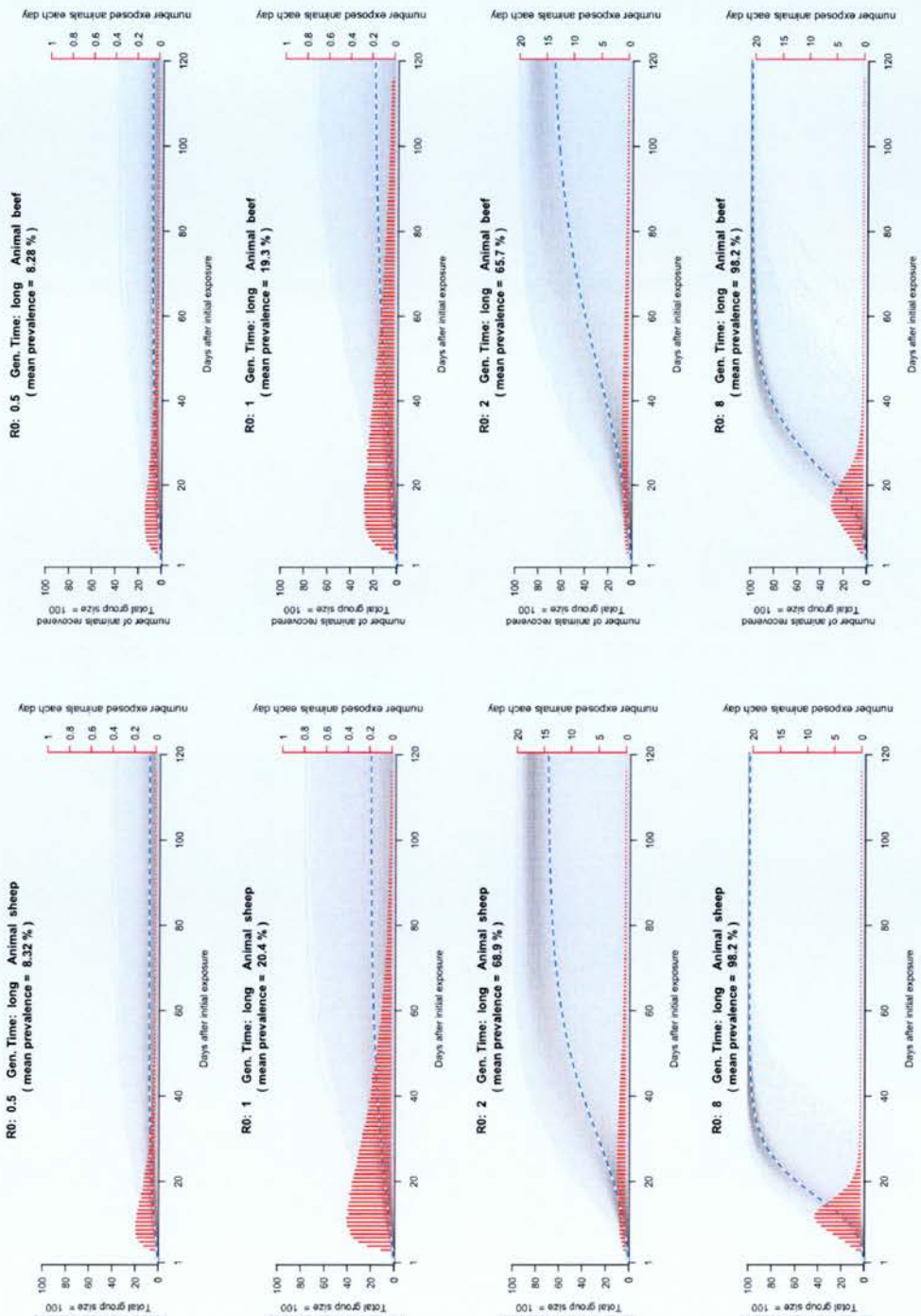


Figure 6.11: Output from the SEIR model for scenarios 1–8 with long incubation and infectious periods showing distribution number of recovered animals against time (grey density plot), mean output from model of recovered animals (blue/white dashed line) and daily new number of exposed animals (red bars). The plots are the result of 10,000 simulations in a group of 100 animals.

6.4.3 Parameters to describe clinical profile of disease

There are four sets of parameters needed for the clinical detection model:

- A description of the relative likelihood of detection from infection to resolution of clinical signs
- An estimate of the maximum probability that disease would be detected in an infected animal if the animal is observed
- An estimate of the probability that an individual animal is observed if the group is observed
- An estimate of the probability that the group of animals is observed on a random day

The clinical profile is composed of a symptomless incubation period (see figure 6.1) followed by a period of clinical signs.

6.4.3.1 Incubation Period

Alexandersen et al. [2003] suggest that the within farm incubation period is generally 2–14 days but that it can be as low as 24 hours in pigs or in high challenge scenarios. They record the *typical* incubation period in sheep and cows as 2–6 days. In experimental conditions they report the mean incubation period as 3.5 days for direct cow to cow infection and 2 days for intensive sheep to sheep infection. For the model I have assumed a deterministic incubation period of 2 days for sheep and 3 days for cattle (the model structure does not allow fractional incubation periods)

6.4.3.2 Clinical Signs

The time profile of clinical signs was informed by literature based data from [Alexandersen et al., 2003; AVIS Consortium, 2002; Kitching, 2002; Kitching and Hughes, 2002] and discussion with a foot-and-mouth disease expert. The result was table 6.7 giving an ordinal clinical detection profile shown in table 6.8.

Day of clinical signs	Clinical score	Clinical signs
1	1	Pyrexia Hot painful feet Reluctance to stand or walk Inappetence Reduced milk production
2	2–3	Nasal discharge Early vesicles in mouth Increased salivation and drooling
3	3–4	Mucopurulent nasal discharge Further vesicles formed in mouth Vesicles on coronary band, Inter-digital cleft and bulb of heel Teeth grinding due to pain Acute lameness
3-4	4	Ruptured vesicles
4–5	4	Lesions show fibrin deposition Epithelial re-growth
5–7	3–2	Scar tissue formation and healing Normal body temperature restored
7–14	2–1	Healing of lesions

Table 6.7: Approximate clinical progression of foot-and-mouth disease in cattle by day since appearance of clinical signs. An ordinal clinical detection score is estimated from these signs and adapted for sheep with discussion with an epidemiologist and a clinician (Mark Bronsvoort and David Black (Personal communication 2008)).

Day	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Cattle	0	1	1	2.5	3.5	4	4	3	2.5	2	2	1	1	1	1	1	1	1
Sheep	1	2.5	3.5	4	4	2	1	1	0	0	0	0	0	0	0	0	0	0

Table 6.8: Daily clinical detection score based on table 6.7. (Day is days since infection)

Animal	Probability	Estimate	Notes
Sheep	Observe flock	0.14	Weekly observation on average
	Observe individual	0.40	Spatial dispersion
	Detect if diseased	0.10	Subtle clinical signs
Beef	Observe herd	0.66	Every 1.5 days
	Observe individual	0.90	Majority of herd cluster
	Detect if diseased	0.80	Some animals observed at distance
Dairy	Observe herd	1.00	Twice daily milking
	Observe individual	1.00	Milking animals
	Detect if diseased	0.95	Obvious signs and close observation

Table 6.9: Estimates of observation parameters used to inform the clinical detection model.

6.4.3.3 Remaining clinical detection parameters

These estimates were based on the interviews summarised in appendix 6.8. The probabilities and justifications are shown in table 6.9

6.4.4 cELISA Sensitivity estimates over time

The estimated parameters for the four parameter logistic model of cELISA test response are shown in tables 6.10 and 6.11 for cattle and sheep respectively. These parameters are used to simulate cELISA percentage inhibition measurements for cattle and sheep as shown in figures 6.12(a) and 6.12(b) together with the observed cELISA results from Paiba et al. [2004].

	Mean	SD
A	11.09	3.80
B	97.16	0.62
C	10.76	33.53
D	7.39	0.270
E	−6.01	0.59
F	0.16	0.14
G	0.0011	0.0042

Table 6.10: Estimates from JAGS model of parameters for logistic model of cELISA test response in cattle. Results are from study where animals were infected by contact with inoculated animals.

	Mean	SD
A	5.9389	3.1526
B	96.7790	0.7705
C	12.1681	34.7948
D	5.9871	0.3567
E	-4.0273	0.6161
F	-1.0064	0.2246
G	0.1088	0.0172

Table 6.11: Estimates from JAGS model of parameters for 4 parameter logistic model of cELISA test response in sheep. Results are from study where animals were infected by contact with inoculated animals.

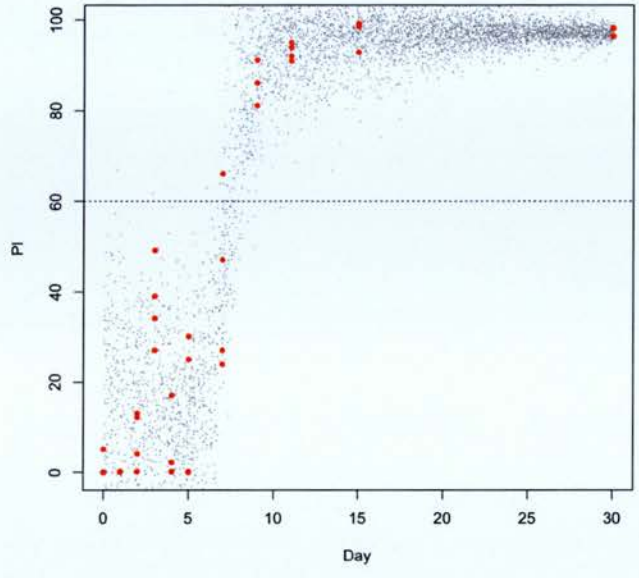
Day	Mean	Empirical
1	0.011	0.00
2	0.006	0.00
3	0.003	0.00
4	0.001	0.00
5	0.001	0.00
6	0.002	0.00
7	0.065	0.25
8	0.814	
9	0.993	1.00
10	0.999	

Table 6.12: Estimates of cELISA test sensitivity in cattle with time using data from [Paiba et al., 2004] estimated with JAGS Bayesian model. Estimates are shown by day until approximate convergence with 100% sensitivity (estimated to 30 days). Results are from a study where animals were infected by contact with inoculated animals. Simple proportion empirical estimate from the raw Paiba et al. [2004] data are also shown for indirect inoculation results. The empirical results are the simple proportion of test animals whose percentage inhibition results were above cut-off value of 60%.

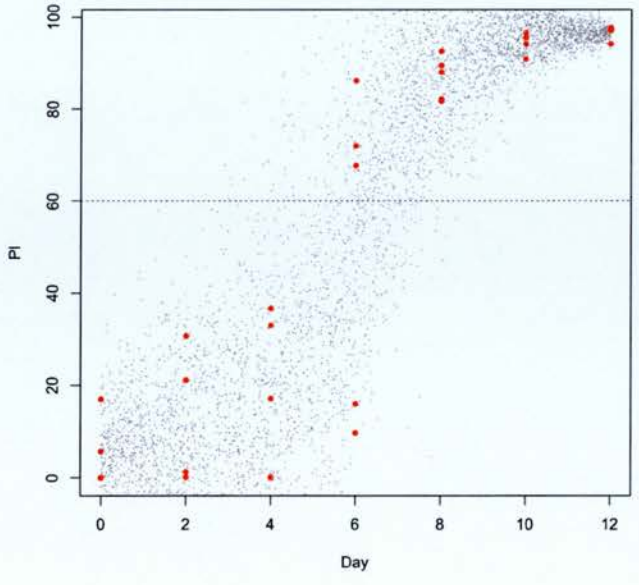
Day	Mean	Empirical
1	0.001	
2	0.003	0.00
3	0.012	
4	0.032	0.00
5	0.085	
6	0.348	0.60
7	0.831	1.00
8	0.981	
9	1.000	
10	1.000	1.00
11	1.000	
12	1.000	1.00

Table 6.13: Estimates of cELISA test sensitivity in sheep with time using data from [Paiba et al., 2004] estimated with JAGS Bayesian model. Estimates are shown by day until approximate convergence with 100% sensitivity (estimated to 30 days). Results are from a study where animals were infected by contact with inoculated animals. Simple proportion empirical estimate from the raw Paiba et al. [2004] data are also shown for indirect inoculation results. The empirical results are the simple proportion of test animals whose percentage inhibition results were above cut-off value of 60%.

The resulting estimates of cELISA diagnostic sensitivity versus time after exposure are shown in figure 6.13 and in tables 6.12 and 6.13 where they are shown with corresponding simple estimates from the raw data. Sensitivity in cattle and sheep is virtually zero until 5 days post exposure in sheep and 7 days post exposure in cattle. It then increases rapidly to effectively 100 % by 8-9 days after exposure.



(a) Cattle



(b) Sheep

Figure 6.12: Observed percentage inhibition from cELISA [Paiba et al., 2004] (red dots) and simulated values (jittered black points) for cattle and sheep with indirect foot-and-mouth disease virus inoculation. Standard cut-off shown with horizontal dotted line.

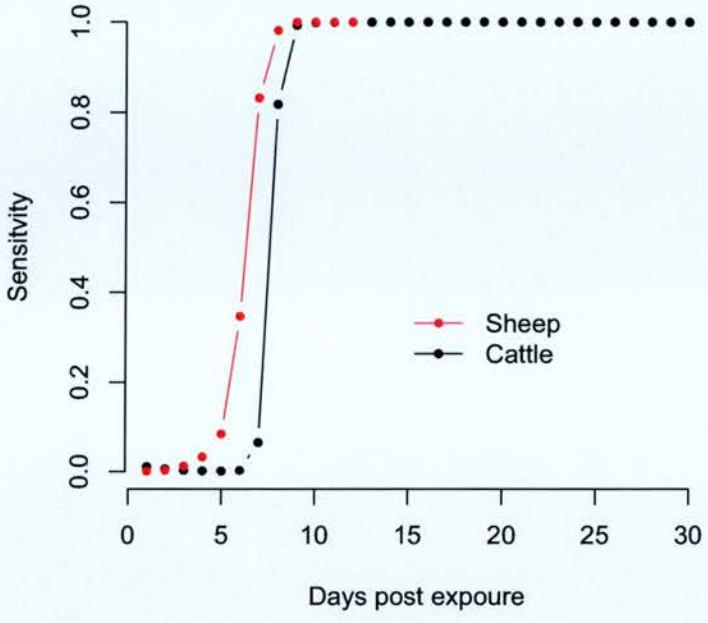


Figure 6.13: Sensitivity estimates of cELISA [Paiba et al., 2004] versus time using JAGS model for *indirect* inoculation

6.4.5 Simulation Results

Figure 6.14 shows illustrative output of the model estimating posterior probability of disease for a selection of scenarios. This decreases from a starting, prior probability, as time passes with repeated clinical surveillance given that no clinical disease is detected. In these results there is a lag in the probability decrease due to the incubation period of foot-and-mouth disease during which no clinical signs will be detectable. Once clinical signs of foot-and-mouth disease would have appeared there is a relatively rapid decrease in posterior probability of disease until the probability approaches an asymptote. The asymptotic probability is reached when the outbreak of disease within the herd or flock has extinguished and all clinical signs have resolved. In the case of sheep and beef animals the asymptote is reached relatively slowly as the observation pattern and probability of detection of clinical signs in a diseased individual are relatively low so, on average, only a small amount of information is gained each day. In a dairy herd the close and frequent observation of the animals and rapid appearance of clinical signs means that in a diseased herd clinical signs will be rapidly detected so most information is gained in the first few days of observation. Serological surveillance at the end of a given period of clinical surveillance decreased the time to reach a particular posterior probability of disease and decreased the asymptotic posterior probability of disease in sheep and beef animals. In dairy animals there were no significant sensitivity gains by using serological surveillance.

Figures 6.15 and 6.16 show the asymptotic posterior probabilities for clinical surveillance alone and clinical surveillance followed by a serological survey at the end of the clinical surveillance period (i.e. when clinical surveillance has reached its asymptote). As would be expected the posterior probability is lower with higher final prevalence (higher R_0) epidemics in beef and sheep groups. In dairy animals clinical observation is highly effective in all epidemiological scenarios almost instantaneously giving a near zero posterior probability of disease given no clinical signs are observed. In contrast in sheep and beef groups serological surveillance gives a marked reduction in posterior probability of disease, given all animals test negative. Serological surveillance when the whole group is sampled is particularly effective. Performance of sample based serology compared to whole herd serology decreases as an adjunct to clinical surveillance in sparse outbreaks. (i.e. gives a higher posterior probability of disease in sparse outbreak scenarios)

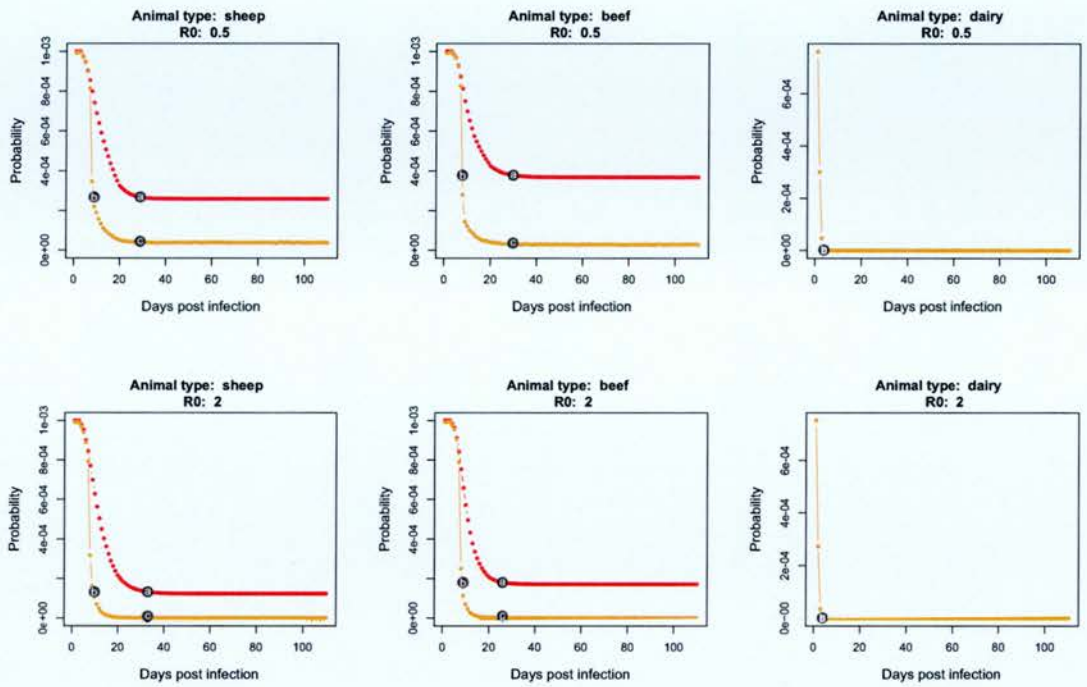


Figure 6.14: Selected output from model showing estimated posterior probability of disease decreasing with time from potential exposure give no clinical signs are detected (red line). The instantaneous benefits of serological surveillance are shown with the orange line representing the posterior probability of disease with time given no clinical signs of disease and a negative serology survey on the date in question. Point (a) is an estimated probability and date at which clinical surveillance offers no significant reduction in probability of disease within the tolerance parameter. Point (b) show that this probability may be achieved (often earlier) by combining serological surveillance and clinical surveillance. Point (c) shows the improvement in confidence of disease freedom available by doing serological surveillance at the date at which clinical surveillance ceases to have benefits. In these examples the serological surveillance was based on a random sample of the animals in the group (sample size designed using FreeCalc) and the prior probability of disease was set to 0.0001. In the case of dairy cattle the clinical detection is sufficient good to give no gain when serological surveillance is added — the orange and red lines are coincident.

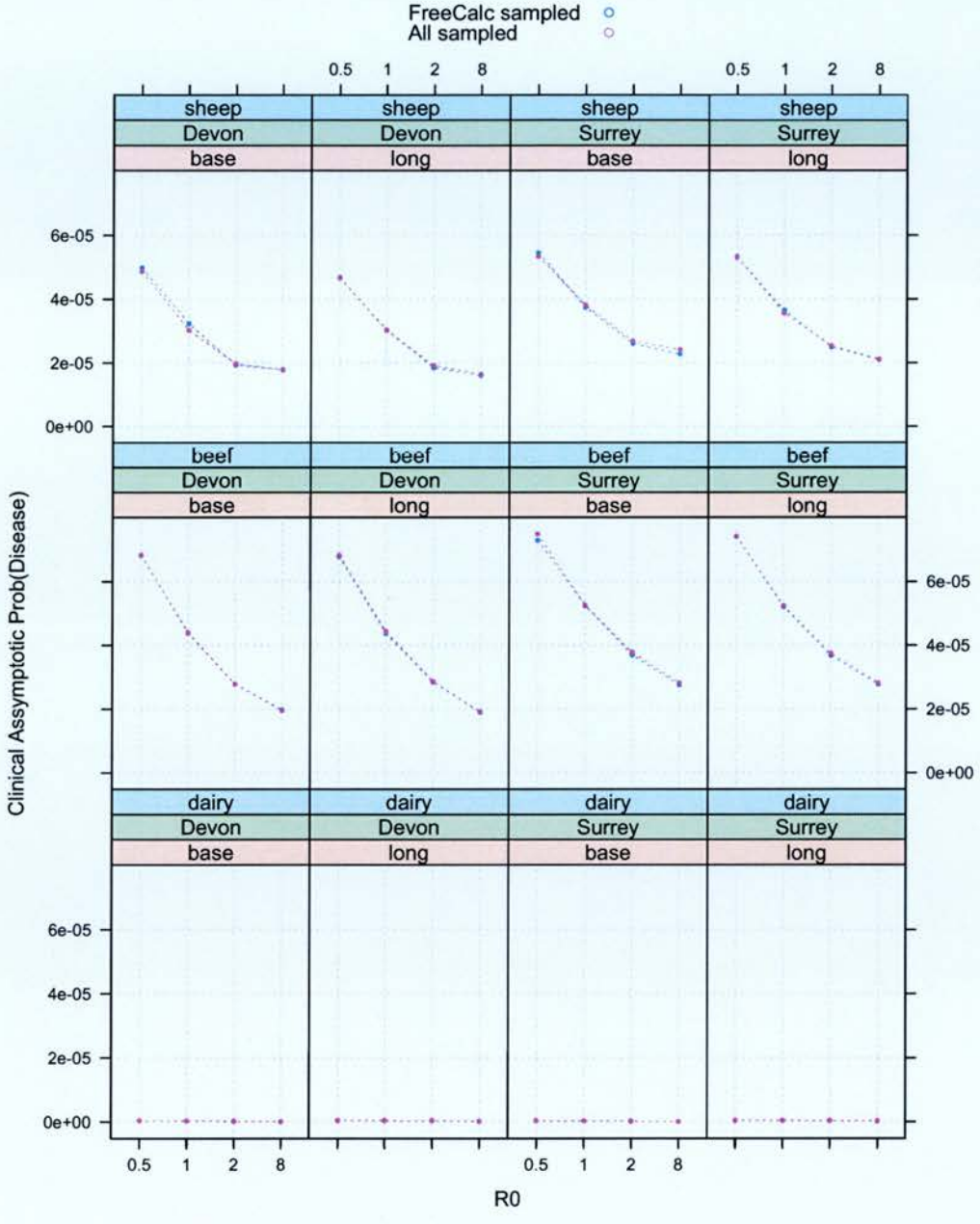


Figure 6.15: Asymptotic posterior probability of disease following repeated clinical surveillance. Shown for each R_0, generation time (base or long), for each husbandry system and county.

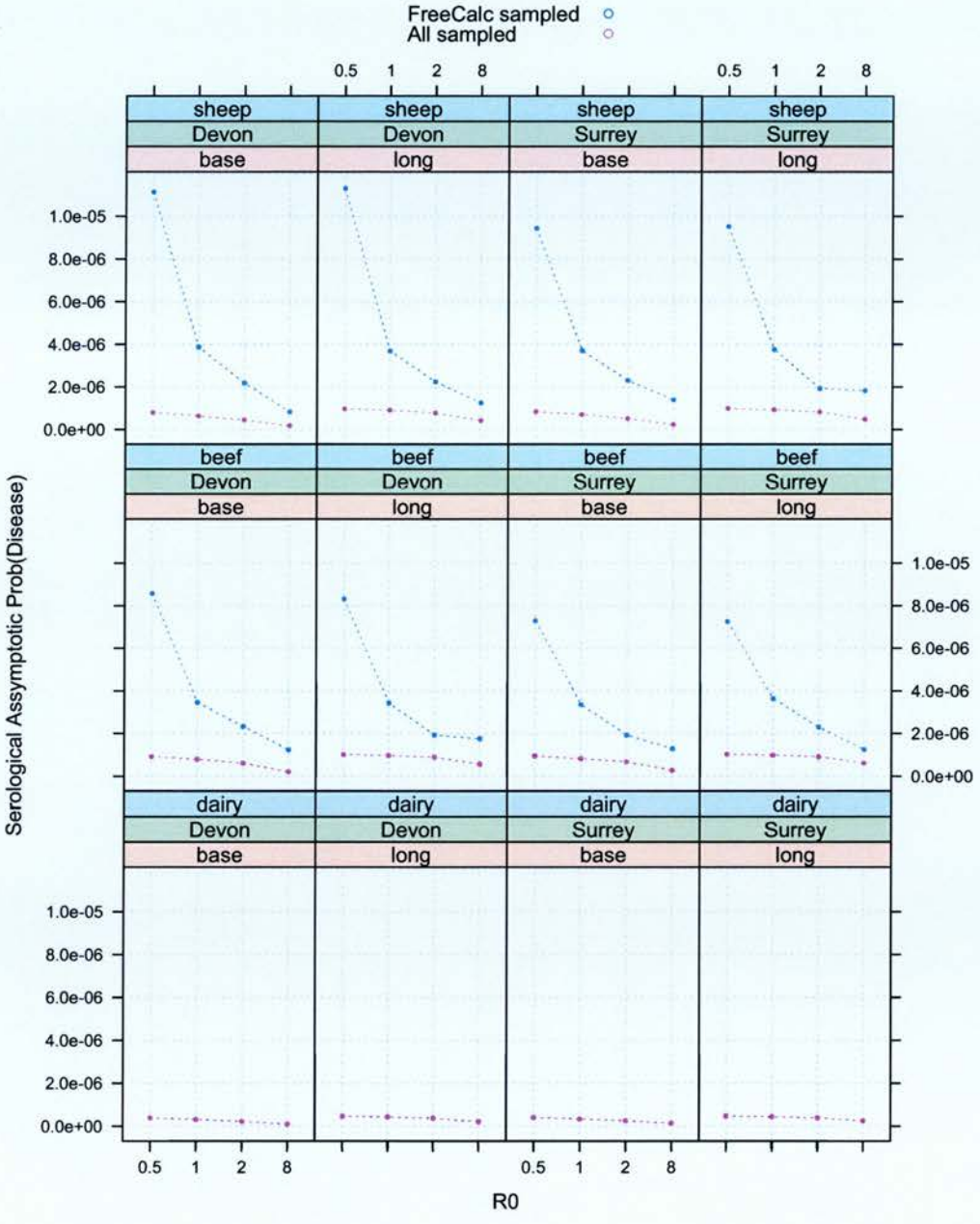


Figure 6.16: Asymptotic posterior probability of disease following repeated clinical surveillance with serological surveillance at the end of the period. Shown for each R_0, generation time (base or long), for each husbandry system and county. Results are shown for complete animal group sampling and FreeCalc based sampling.

Figures 6.18 and 6.17 show summary results of simulations over a range of epidemiological scenarios, animal types, sampling strategies and demographics. These results were robust to the tolerance used to determine the asymptote for posterior probability as discussed in section 6.3.5 so are reported using an asymptote detection tolerance of 1%.

Figure 6.17 shows the relative effect of whole herd/flock sampling versus random sampling where the random sampling is designed to meet a pre-determined within herd/flock sensitivity criteria (in this case 95% sensitivity assuming 5% within-group prevalence). In the high R_0 (high final prevalence scenarios) there are no marked benefits to whole herd sampling. However as final prevalence decreases (still remaining above 5% on average) there are marked benefits (i.e. reductions in posterior disease probability) to be gained by sampling the entire group of animals for sero-surveillance. These gains are slightly greater in sheep flocks in the simulations using Devon based demography than Surrey based demography. The gains are a consequence of the improved detection of low prevalence disease with full group sampling. Full group sampling using a sensitive test (like the cELISA test used in this simulation study [Paiba et al., 2004]) is almost 100% sensitive at a group level as shown in figure 6.16 which shows the *absolute* asymptotic performance when serology is performed additionally to clinical surveillance — probabilities of disease being reduced from a prior of 0.0002 to a posterior of 1×10^{-6}.

The figure (6.17) shows the posterior probability of disease reduction obtained by performing serological surveillance at the end of a period of clinical surveillance. The absolute value of this reduction will be dependent on the prior probability of disease chosen. However, the prior probability of disease was found only to influence the scale of the probability reduction metrics so the results are discussed using the single prior probability of disease of 0.0002. The posterior probability reduction by adding sero-surveillance decreased in sheep and beef groups across the eight epidemiological scenarios as R_0 and final prevalence increased (the final prevalences are shown in table 6.6).

Figure 6.18 shows the number of days saved by adding end of period serological surveillance. Assuming the asymptotic posterior probability of clinical surveillance is a target performance, this probability can be obtained much earlier in beef and sheep groups by adding serological surveillance after a shorter period of clinical surveillance. No significant time is saved in dairy cattle as clinical surveillance achieves low posterior

probabilities very quickly due to twice daily observation and high detection probabilities. In beef (non-dairy) groups the saving is approximately 15–30 days (assuming R_0 is 2.0) and in sheep flocks approximately 20–33 (assuming R_0 is 0.5). The greater saving in sheep flocks is a consequence of the poorer clinical detection in these groups. Over the range of R_0 scenarios in sheep flocks the time saved increases, at first, with decreasing final prevalence then decreases over the last two to three scenarios. This effect is consistent over the range of prior probabilities of disease, tolerances of asymptote detection and demographic scenarios. This apparently paradoxical effect is a consequence of the measure used which is, in full, the number of days saved by using serology to achieve the asymptotic clinical only probability. In the $R_0 = 0.5$ scenario the asymptotic probability with clinical surveillance is markedly higher than in the other scenarios. The asymptote is reached relatively quickly so although serology is still beneficial in these scenarios the apparent time saved is less.

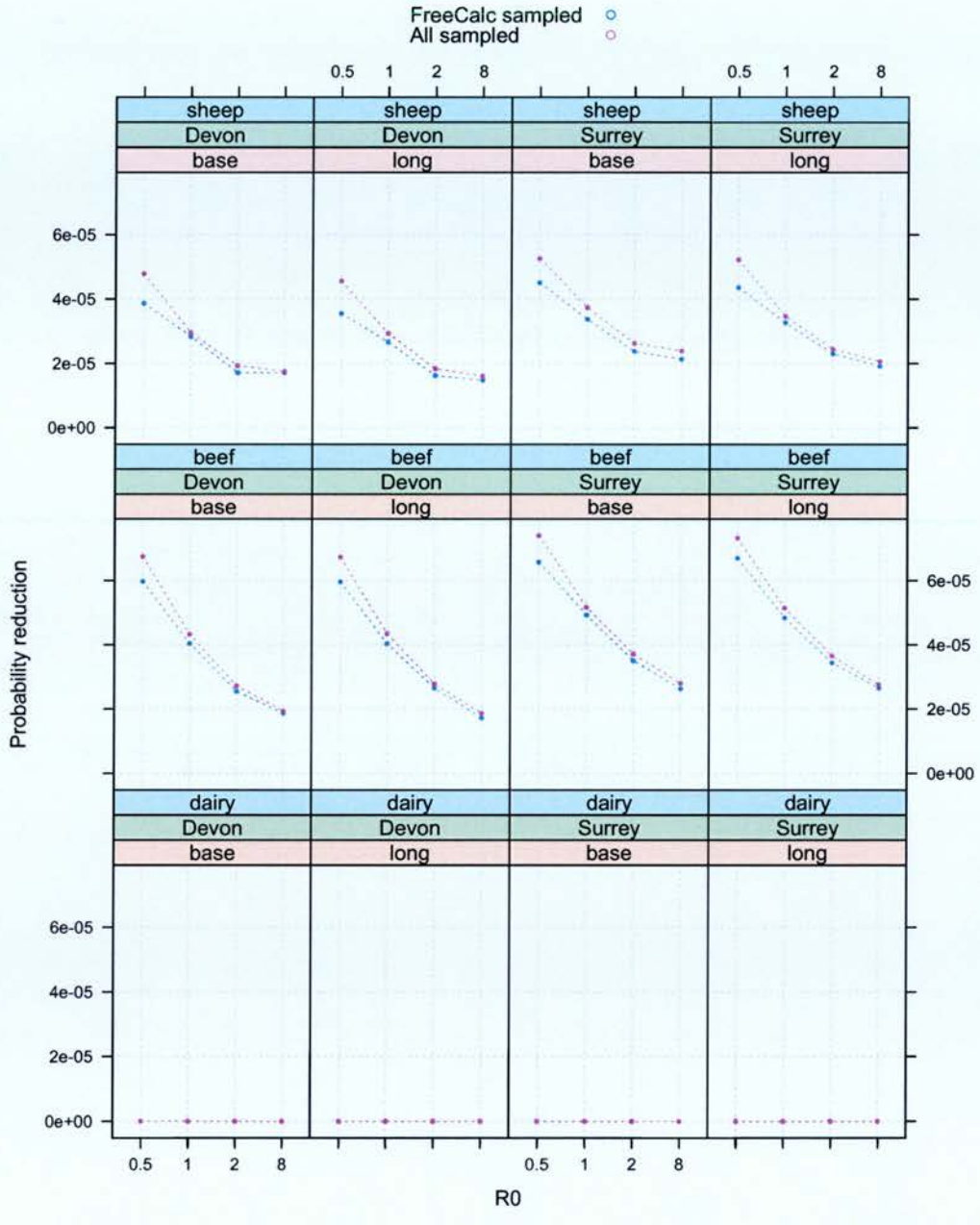


Figure 6.17: Estimated estimated reduction in posterior probability of disease in an animal group by doing sero-surveillance as an adjunct to day-by-day clinical surveillance assessed at the date after exposure when clinical surveillance ceases to give significant additional information. Shown for each $R0$, generation time (base or long), for each husbandry system and county. Results are shown for complete animal group sampling and FreeCalc based sampling.

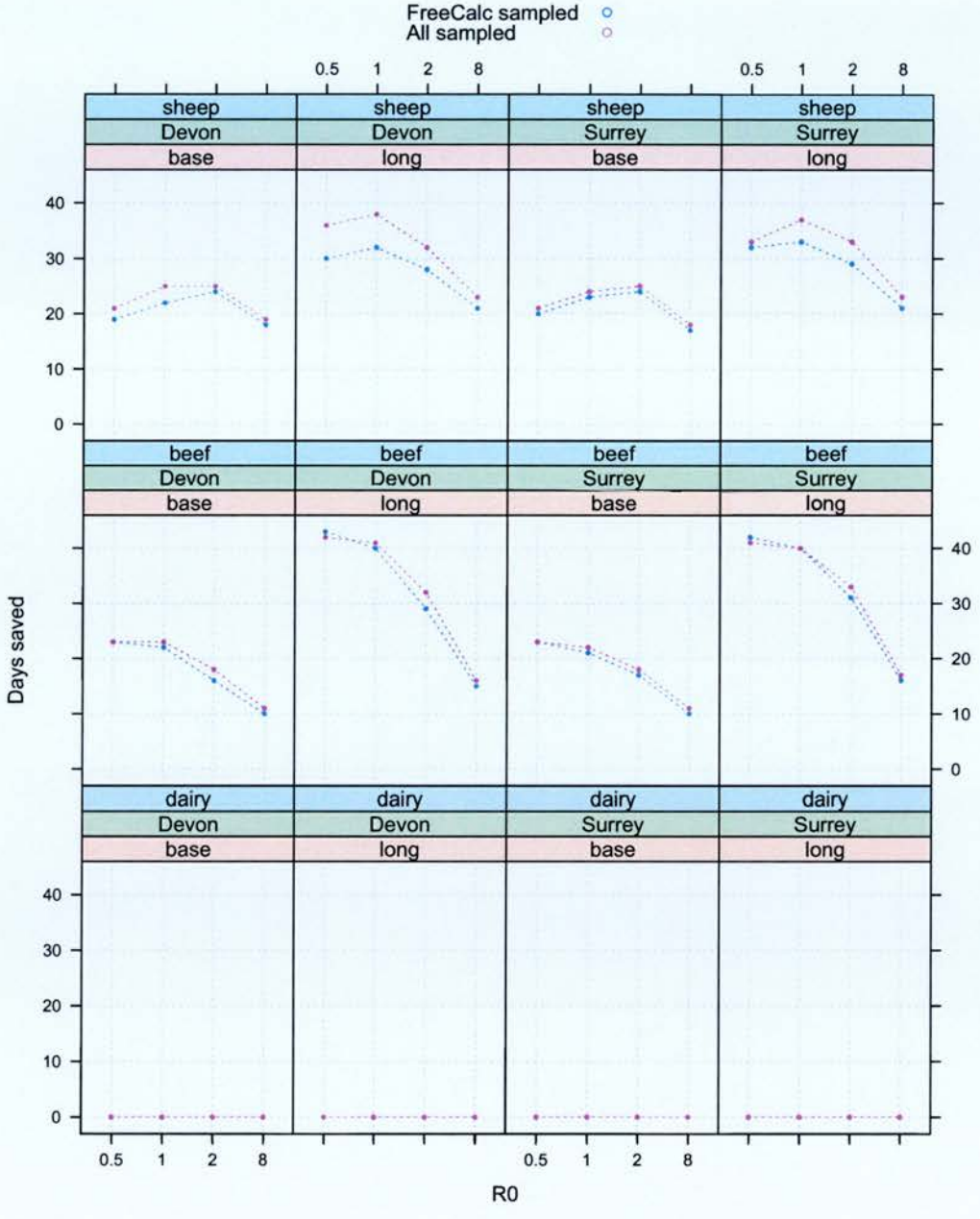


Figure 6.18: Estimated number of days saved to achieve a similar posterior probability of disease in an animal group by doing sero-surveillance as an adjunct to day-by-day clinical surveillance. Shown for each $R0$, generation time (base or long), for each husbandry system and county. Results are shown for complete animal group sampling and FreeCalc based sampling.

6.5 Discussion

6.5.1 General discussion

This study has estimated, for a set of scenarios, the daily probability that a group of animals is diseased given a starting belief and absence of clinical signs during continuous clinical surveillance. The decrease in the probability, from an assumed time of exposure, is rapid in the case of milking dairy cattle and markedly slower in the case of non-dairy cattle and sheep. Traditionally, serology has been considered more important in sheep than in cattle but these results emphasise that consideration should be given to the potentially less effective detection in beef animals. The findings in the UK 2007 foot-and-mouth disease outbreak where, at detection, clinical signs on the index beef farm were thought to be up to 10 days old [Ryan et al., 2008] reinforce this. The ultimate probability of disease, given no observed signs, is lower in dairy animals than in other systems. Serological surveillance by partial or complete blood sampling of the animal group may be performed at the end of a period of clinical surveillance. In dairy animals this surveillance does not significantly improve either the ultimate confidence of disease freedom or the speed with which that confidence is obtained. In non-dairy cattle and sheep, serological surveillance will give marked improvements in confidence of disease freedom and/or speed by which a required confidence is achieved. The same confidence in disease status of a group may be reached approximately 2–4 weeks sooner using sero-surveillance. The relative benefits of serology are greatest in sparse outbreaks with poor clinical observation. Such settings will also increasingly reward more thorough serological surveillance, with sampling proportions greater than the traditional design requirements.

The decision to undertake serological surveillance in addition to clinical surveillance may be solely driven by regulatory and political pressures; previously without serological surveillance an affected country could not regain the status of disease-freedom and resume international meat and livestock trade. OIE regulations have somewhat relaxed and now require portfolios of evidence rather than prescriptive survey designs [OIE, 2009b]. It is unlikely that serological surveillance would disappear from such portfolios but it is efficient to apply serology where it is most beneficial, i.e. where clinical surveillance is least likely to provide sufficient evidence. This has always been the case to an extent, this study helps to quantify these benefits.

The simulation model has only dealt with a single group of animals where it is assumed that a number of animals are instantaneously exposed at the start of the simulation period. In a foot-and-mouth disease outbreak post epidemic surveillance will normally have to consider the more epidemiologically and logistically complex problem of a heterogeneous set of animal holdings which are all at some risk of exposure over a wide time window. On a farm to farm basis serology will always have the benefit, in non-dairy animals, of giving a marked improvement in confidence of disease freedom. The benefit that serology reduces the time to declaring a group of animals disease-free is harder to quantify in a heterogeneous multiple farm setting. If it is assumed that all farms are exposed at the start of the post epidemic period the timing improvements of serology will be similar to those of the single groups used in this study. However foot-and-mouth disease virus is a challenging pathogen. In the UK 1967 epidemic several farms experienced outbreaks up to six weeks after infection was apparently under control with no infected individuals [HMSO, 1969, page 41]. In the UK Surrey 2007 outbreak environmental contamination both initiated the outbreak and possibly resulted in a delayed secondary outbreak some distance from the index infected premises [Anderson, 2008]. This was assumed to be due to environmental reservoirs of foot-and-mouth disease virus. Uncertainty about the environmental reservoirs [Alexandersen et al., 2003] and 'silent-spread' [Blanco et al., 2002] of foot-and-mouth disease will add uncertainty to the precise benefits of serological surveillance. These issues are in part managed by the regulatory guidelines that require serological surveillance to normally commence only after a suitable delay (previously 2 months after the last notified case of foot-and-mouth disease). Whilst this wait period makes it more difficult to precisely value the time benefits of serology the time benefits are still potentially gained — if environmental recrudescence effectively exposed a group of animals near to the end of a 60 day wait period a decision about that group reinforced with prompt serology potentially has much greater confidence than without it i.e. it is equivalent to waiting 60 days then an additional long period of clinical surveillance alone.

6.5.1.1 Assumptions and limitation

In this analysis I have chosen to only consider surveillance in cattle and sheep. Post epidemic surveillance in both of the recent UK foot-and-mouth disease epidemics was mainly an issue in these species. Whilst a poorly managed pig unit was the index farm of the 2001 epidemic, the disease went on to be largely one of sheep and cows [Gibbens

et al., 2001]. In areas of intensive pig production pig unit bio-security is high and epidemiology of disease rapid so clinical detection would be likely to be effective and rapid. The high value of large scale pig production units combined with the potential for high volume virus excretion from diseased pigs means that strategic vaccination combined with rapid culling of infected premises of pig units may be a possible response in a future epidemic if the strain and location place the pig population at significant risk [TAIEX, 2007].

This study has not considered surveillance in a vaccinated population. Vaccination will have marked influences on the precise benefits of clinical and serological surveillance. Disease in a vaccinated or partially vaccinated population may have markedly different epidemiological characteristics [Orsel et al., 2007a]. Vaccination can suppress clinical signs but may still permit virus replication and excretion [Paton et al., 2006]. Also serological detection in vaccinated animals will require different non-structural protein based diagnostic techniques [Sorensen et al., 1998]. These tests are generally less sensitive than the c-ELISA test used in un-vaccinated stock. Hence while clinical and serological surveillance will both still have benefits after vaccination-based control the relative quantitative value of these benefits may be quite different. More generally, serology may have additional benefits, even in unvaccinated populations, if a foot-and-mouth disease strain produces particularly subtle or absent clinical signs in a species as clinical detection will be relatively less effective.

For modelling and descriptive simplicity I have artificially categorised the study populations into dairy, non-dairy (beef) and sheep. The epidemiological and surveillance models estimate the surveillance performance in homogenous units. In reality a large number of farms are mixed (see tables 6.1 and 6.2) and certainly all dairy farms will have some degree of mixing between milking animals and replacement heifers, dry (non-milking) cows, calves and bulls. On these farms there may well be epidemiological mixing such that disease in one class of animal may spread to others. In this setting there may be a local sentinel effect whereby, for example, a few dairy animals kept in proximity to a large number of lowland sheep become infected with foot-and-mouth disease after the sheep flock but will be detected first due to more pronounced clinical signs and more attentive observation. These effects will be difficult to estimate as on farm mixing will be highly variable from farm to farm. The overall effect would be likely to improve clinical detection performance when sheep and cattle are mixed.

On farm management behaviours change markedly throughout the year in the UK cattle and sheep sectors. Grazing of animals is driven by grass production with, for example in some systems, beef stock being moved from more remote pastures to indoors housing for the winter months. The sheep industry is not homogenous with lowland and highland sectors and extensive movement between pasture types throughout the farming year. These movements mean that clinical observation of each animal type will be highly contingent on the system and time of year when an epidemic strikes. It is possible that these influences are partially self compensating; from an onwards risk perspective, although clinical surveillance in extensively grazed hill sheep will be very poor the scale of epidemics and onwards risk of infection to other systems will be low.

The serological detection model required estimates of day by day sensitivity of the cELISA diagnostic test. The smoothing model used gives a small but non-zero sensitivity for the diagnostic test in the first few days after exposure. The results reported use this model, setting the first two day's sensitivity to zero did not make a significant change to the results. Also, in the model, the test was considered to have a specificity of 100% i.e. there was a zero probability of classifying a herd as diseased if none of the sampled animals were diseased. In reality all diagnostic tests will produce some false positive results, whilst these results in themselves are diagnostically wrong they do have the paradoxical side-effect of increasing herd level sensitivity of a serological survey; i.e. a herd that is diseased may be flagged as diseased despite no animals with disease being tested. The same is particularly true of clinical surveillance where there are anecdotal reports of false positive results during the control phases of the UK 2001 epidemic. These results if thoroughly followed up with expert investigation and further testing will increase the performance of the surveillance system but at potential considerable cost due to consumption of resources in the investigation potentially uninfected farms.

When herds or flocks are blood sampled for serological surveillance the procedure will inevitably involve close contact with the livestock. Hence there is a possibility of increased rates of *clinical* detection during *serological* surveillance. For the purposes of analysis I have assumed this effect is negligible. In practice it may slightly increase the benefits of serological surveillance when this is considered as a whole procedure involving sampling and serological testing.

In this study the detection systems are simply classified into clinical and serological. Clinical detection will include visual and physical examination with perhaps the sophistication of milk production measurements. Current research elsewhere is exploring other modalities of detection. As these are introduced it would probably be necessary to reclassify the approach used in my study into continuous and one-off surveillance. For example the use of thermal imaging cameras [Rainwater-Lovett et al., 2008] allows veterinary officers to visit a farm and identify animals with higher that normal body/peripheral temperature areas as an early indication of clinical disease. Thurmond and Perez [2006] model the use of bulk milk tank based detection systems with pooled analysis of daily samples from dairy farms. Other longer term ideas include the use of low cost embedded or ruminal temperature measuring probes that provide telemetry back to the farm reporting an almost continuous record of body temperature. The former thermal imaging observation may be used in a functionally similar way to a serological survey — it takes place on a single occasion and provides a single opportunity for detection. The temperature telemetry approach is more akin to clinical surveillance where, with suitable analysis algorithms, it provides a continuous opportunity for a disease alert trigger.

6.5.1.2 Concluding remarks

The simulation model has suggested, within the outbreak scenarios used, that serology only has significant benefits in sheep and non-dairy settings. Serology at the end of a period of clinical surveillance can be used to either markedly increase the sensitivity of disease detection in the group or to bring detection (to a predetermined threshold) forward by perhaps 15–33 days. Serology is relatively more beneficial in sparse epidemics with a low R_0 and low final prevalence, such as upland sheep flocks, particularly if applied to whole groups rather than small random samples. It also detects historical infection more effectively than clinical surveillance, antibodies persist long after foot-and-mouth disease lesions have healed.

6.6 Appendix 1 — Formulation of JAGS model to estimate cELISA test sensitivity

$$\left.\begin{aligned} p_i &\sim \text{dnorm}(\mu_i, \tau_i) \\ \mu_i &= A + (B - A)/(1 + e^{C\cdot(D - t_i)}) \\ \tau_i &= exp(E + F \cdot t_i + G \cdot t_i^2) \end{aligned}\right\} 1 \leq i \leq N$$

$$\begin{aligned} A &\sim \text{dgamma}(1, 0.01) \\ B &\sim \text{dgamma}(1, 0.01) \\ C &\sim \text{dgamma}(1, 0.01) \\ D &\sim \text{dgamma}(1, 0.01) \\ E &\sim \text{dnorm}(0, 0.01) \\ F &\sim \text{dnorm}(0, 0.01) \\ G &\sim \text{dnorm}(0, 0.01) \end{aligned}$$

$$\left.\begin{aligned} \mu_{1,j} &= A + (B - A)/(1 + e^{C\cdot(D - j)}) \\ \tau_{1,j} &= exp(E + F \cdot j + G \cdot j^2) \\ \text{psim}_{1,j} &\sim \text{dnorm}(\mu_{1,j}, \tau_{1,j}) \\ \text{psim}_{1^\star,j} &= \text{psim}_{1,j} - \text{cutoff} \\ \text{se}_j &= \text{step}(\text{psim}_{1^\star,j}) \end{aligned}\right\} 1 \leq j \leq \text{Nt}$$

6.7 Appendix 2 — Transmission data from literature

1 — Young calves Karin Orsel [Orsel et al., 2005]

8-10 week old Holstein calves housed in free mingling groups

Challenged with O/NET/2001 [Bouma et al., 2004]

Two each out of generally four calves inoculated with virus

Four fold increase in VN titre or clinical signs or PCR or NS-ELISA was diagnostic test for foot-and-mouth disease

Experiment ended at 30 days post infection

Used final size analysis to estimate R_0 by MLE

Vaccinated calves — R_v 0.18 (0.01–1.2)

Unvaccinated calves — R_0 2.52 (1.13–52.1)

2 — Dairy cows Karin Orsel [Orsel et al., 2007a]

Two groups of 10 Holstein-Fresian dairy cows (2,3, and 4th parity 50-120 days post partum) (in unvaccinated experiment)

Random milking routine and normal bio-security for milkers

Challenged with O/NET/2001 [Bouma et al., 2004]

5 cows inoculated in each group of ten

Experiment lasted 31 days

Four fold increase in VN titre considered diagnostic

Used final size analysis to estimate R_0 by MLE

Odd estimates from analysis - see paper - dependent on assumption of infectiousness of inoculates

All contacts became infected after exposure to the inoculates in unvaccinated animals.

No transmission occurred to vaccinates.

R_v estimates were 0 for vaccinates and ∞

for unvaccinated cows.

3 — Lambs Karin orsel [Orsel et al., 2007b]

52 weaned lambs — 10 weeks old

Assigned to 13 groups of 4. Two lambs per four were foot-and-mouth disease inoculated.

Four fold VN titre increased used as diagnosis.

Experiment ran for 30 days.

GLM as per Klinkenberg [Klinkenberg et al., 2002] to estimated β

Also estimated infectious period

β in unvaccinated lambs 0.105 per day (0.044 – 0.253)

T estimated as 21.1 (10.6 – 42.1)

Hence R as 2.22 (0.46 – 4.33)

In vaccinated lambs limited data but final size estimate of R_v as 1.14 (0.3 – 3.3)

4 — 1000 dairy cow simulation model [Carpenter et al., 2004]

5 — Sheep [Streftaris and Gibson, 2004] Two experiments of 32 sheep each.

Sheep in groups of 8

Naturally infected sheep used for parameter estimates
Bayesian estimates of parameters of SEIR model
Experiment 1
β 0.024 (0.005–0.048)
latent period 1.599 (0.440–1.507)
Experiment 2
β 0.020 (0.004–0.063)
latent period 1.937 (0.836–3.403)
No record of infectious period estimates

6 — Cattle [Bankowski et al., 2008] Four animal experiments using Holstein cattle. Donor and recipient animals used to estimate latent period and infectious period.
Latent period:
Overlap method: Median 53.33 hours (95% interval 42.35–71.86) Entire period method: Median 68.75 hours (95% interval 58.15–86.80) Infectious period: 106.99 hours (95% interval 40.34–232.41)

7 — Cattle and small ruminants [Mardones et al., 2008] Systematic review and modelling of disease stage duration.

Incubation period:
Cattle: 6.1 days
Small ruminants: 4.7 days

Infectious period:
Cattle: 5.9 days
Small ruminants: 2.8 days

Latent period:
Cattle: 1.8 days
Small ruminants: 2.0 days

6.8 Appendix 3 — Summary of interviews

Interviews were conducted with a sheep clinician/lecturer [Scott, 2008], a practising sheep clinician (Black, David. Personal communication August 2008), a lecturer in animal husbandry [Hodgson-Jones, 2008], a farm manager [Aitchison, 2008b] and shepherd [Aitchison, 2008a]. The points made are listed in the following sections. These points represent a summary of the interviews and in some case provide contradictory opinion.

6.8.1 Sheep

1. Sheep farmers are required to examine their stock on a daily basis
2. Hill flocks may actually only be examined with an interval of one week to several months.
3. Sheep are inspected by shepherds during supplement feeding or maintenance activities
4. Inspection is more frequent during lambing, dipping and shearing.
5. When a flock is examined only some animals maybe seen due to wide spatial dispersion.
6. In a post epidemic setting a government vet would probably inspect on weekly basis.
7. During the UK 2001 epidemic veterinary inspections were daily
8. Veterinary officers rarely visited farms after the epidemic unless alerted by the stock keepers or if they were in a high risk area.
9. During the UK 2001 epidemic some hill flocks were gathered and brought down to lowland pastures for increased surveillance however there were concerns that this activity could increase the risk of flock exposure to foot-and-mouth disease and subsequent within-flock transmission.

6.8.2 Cattle

1. Milking dairy cows will be observed at relatively close distance at least twice daily.
2. Dairy farmers will generally know milking cows individually and will be highly likely to identify signs of clinical disease.
3. Current milking machine have the capacity to detect a drop in milk yield and alert the stockman to the problem in an individual cow.
4. Dry-cows (not milking), bulls, calves and heifers (not milking) will be observed at least once a day.
5. Beef cattle are observed less frequently than dairy cattle except during the calving season. They would normally be observed once a day when housed in the winter and every few days when outside during the summer.
6. Suckler cows are observed at least once a day during feeding.
7. In a post epidemic setting the government veterinary officers would rarely visit unless alerted by the stockholder.
8. When a group of beef animals or non-milking dairy animals is observed most (approximately 70%) of the group would be observed.

Chapter 7

Discussion

Demonstration of disease freedom after a disease outbreak or period of endemic disease is important; without it regions and countries will often remain internationally isolated in livestock and animal product trade. However the design of post epidemic surveillance systems is complex, being contingent on the epidemiology of the disease, including its control, the diagnostic test systems available to detect it, the structure and management of the livestock population and the required levels of confidence needed to satisfy stake-holders. Only an instantaneous survey of all livestock in a region with a perfect test can give an unconditional and certain diagnosis of the presence or absence of disease in a population. Real world surveillance systems will always use sampling strategies with probabilistic interpretation of results where diagnostic accuracy of the surveillance system is balanced against the overall cost of implementing the system. In chapter 2, I proposed a targeted or risk based strategy for surveillance system design that aims to reduce the time and resource cost of implementing a surveillance system whilst maintaining a nominated level of diagnostic performance. Selection of sampling units based on mathematical-model, statistical-empirical or expert-based estimates of risk of infection will, if the models are generally predictive, save time and money. Recently, Williams et al. [2009a,b] also consider targeted surveillance using simulation studies and an example scenario of scrapie in sheep. Their work further highlights potential benefits whilst reinforcing the points that targeted surveillance is more suitable for the detection and estimation of prevalence in low *prevalence* settings than for the estimation of prevalence and clustering in higher prevalence settings. Targeted surveillance is dependent on good models of disease. Good models are those that estimate

risk of disease presence in sampled units with relatively low bias and ideally with useful precision. The results in chapter 2 suggest that the benefits of targeted surveillance are, however, relatively robust to model errors though better models will give greater savings with a lower risk of poor survey performance. Targeted surveillance represents an application of mathematical epidemiological modelling in disease control somewhat removed from its previously contentious application in directing efforts at prevention and control but still with potentially great benefits.

Exploring the application and design of animal disease surveillance systems is hindered by the infrequent and episodic occurrence of epidemics and the limited availability of the resulting surveillance designs and data. Access to the ongoing foot-and-mouth disease surveillance work in Thrace, Turkey provided an opportunity in chapters 3 to 5 to estimate the performance of the current approach and to apply a Bayesian methodology to estimate disease epidemiology and test performance from the available survey data-sets.

The simulations in chapter 3 suggest that the classical, two stage, sampling strategy is sufficiently sensitive to detect disease under a range of *design* epidemiological scenarios to satisfy international trade regulations. However two challenges remain with this approach:

Firstly the imperfect specificity of the current, non-structural protein based, diagnostic tests potentially results in a large number of false-positive animal test results. These results then require further confirmatory testing which delays final results, increases costs and reduces the overall surveillance system sensitivity. These effects are made worse in a vaccinated population where repeated vaccination may increase the rate of false positives.

Secondly the estimate of adequate performance assumes that foot-and-mouth disease will be present at a 'design' prevalence of 5% as per previous OIE regulations for demonstration of disease freedom. In a vaccinated population with good vaccination coverage disease prevalence may ultimately be much lower. For example, a model by Arnold et al. [2007] suggests that in a reactively vaccinated population foot-and-mouth disease carrier herds may be present at about 0.2% and would therefore require census based sampling to adequately detect disease. Thus as a region like Thrace moves by vaccination and other control measures towards disease freedom it may be necessary to adjust serological surveillance techniques and use a wider range of evidence

sources to adequately demonstrate freedom from disease. An alternative, as adopted in South America, is to aim to cease vaccination in order to declare disease freedom, permitting more liberal trade and easier demonstration of freedom by sero-surveillance — on the basis that if foot-and-mouth disease were present full-blown outbreaks would be observed.

In chapters 4 and 5 I explored the Thrace surveillance data and made inferences about disease epidemiology and diagnostic test performance. Whilst disease freedom data sets may be used to inform these estimates, the overarching conclusion has to be, particularly with uncertain test performance data and the issue of vaccination related false positives, that estimates of disease prevalence from these data sets are vague with wide credible intervals. This is a consequence of the relatively low information content of the survey for the number of parameters being estimated and the issue of identifiability; it is difficult to jointly estimate, for example, diagnostic test sensitivity and disease prevalence without suitable informative prior information for these parameters.

The Bayesian approach does, though, have definite advantages over the more classical, frequentist interpretation of disease freedom surveys. As a region moves towards disease freedom the probability of disease in the region and group (e.g. village) level prevalence may be tracked to identify trends and improvements. The scenario-tree based methodology of Martin et al. [2007b] explicitly uses Bayes theorem to update previous estimates of the probability of disease with the results from new surveys. Also the OIE phyto-sanitary measures agreement [World Trade Organisation, 2008] requires that international trade of livestock and animal products are only restricted on the basis of scientifically identified asymmetries of risk. Bayesian analysis provides estimates of prevalence which may inform quantitative risk models providing finer grained information than classical disease freedom present/absent models.

Surveillance to gain or regain the status of disease freedom does not depend on serological studies alone; clinical evidence based on observation of a region's livestock is considered an important component but its relative merits are poorly defined. Chapter 6 identified the benefits of serological surveillance relative to clinical surveillance in a range of post-epidemic settings. As with all post-epidemic surveillance the performance of the surveys is contingent on the underlying disease epidemiology. Serological surveillance is at its most useful in production systems with otherwise low rates of observation of stock. Clinical surveillance, as it is present to a degree in all animal

holdings, provides a good surveillance system method when clinical disease is obvious and has a relative advantage over probability-sample based sero-surveillance when disease is sparsely clustered. Serological surveillance will detect subtle disease but only when it is relatively widespread unless sampling is of a high density. The conclusions from chapters 2 and 6 suggest that surveillance, given that suitable models of disease exist, may be designed both in modality (clinical versus serological) and location (using risk models) to give most efficient application of resources to providing prompt and accurate decisions about disease status. Cannon [2009] provides a theoretical overview of different disease detection and estimation methodologies considering efficient use of resources. The article raises the important question that if we are to design surveys efficiently we have to do this with respect to some measure, an objective function, which may be cost, farms visited, samples taken or otherwise. Economically efficient approaches and a careful definition of aims and objective criteria will become increasingly important if cost allocation of animal health moves, as is predicted [Alder, 2008], towards the producer rather than the state.

While serological surveillance is likely to continue to be an important and necessary component of post epidemic surveillance it is important to remember that it is only one piece of a more complex system that collectively helps protect the disease status of a region's livestock. This larger system includes good farming practices that reward a high standard of animal health, appropriate biosecurity measures both in periods of disease freedom and during epidemics and rapid and appropriate responses to new epidemics limiting spread and impact and thus limiting the magnitude of subsequent post-epidemic surveillance tasks. The hazard analysis critical control point (HACCP) based approach of focusing on the high importance/visibility components of animal production pathways may be as important as population based sampling in limiting onwards disease risk in outbreak or epidemic events.

There continue to be further advances in analysis and techniques surrounding post-epidemic surveillance. Past advances have been both theoretical, such as methodologies of sample size calculation and evidence synthesis using scenario trees, and in diagnostics techniques such as the development of non-structural protein based diagnostic tests capable of differentiating between vaccinated and previously infected animals. Future developments may continue along the same threads. Increased computational speeds permit simulation based methods allowing further extensions of risk based survey targeting and Bayesian estimation. New models allow real time estimation of epidemic

parameters and prediction of high risk premises [Jewell et al., 2009a,b]. The development and parameterisation of these complex, individually based, models of disease transmission can be used alongside suitable mechanistic and economic models of the surveillance process to permit countries and regions to regain the status of disease-freedom optimally — balancing costs and delays against the vital confidence in the process needed to reassure stockholders, animal health workers, trading partners and governments.

Bibliography

A. Aitchison. Interview. Easter Bush Veterinary Centre, Midlothian, UK, Aug 2008a.

K. Aitchison. Interview. Easter Bush Veterinary Centre, Midlothian, UK, Aug 2008b.

M. Alder. Busy agenda. *Veterinary Record*, 163(22):639–, 2008. Editorial comment.

S. Alexandersen, Z. Zhang, and A. I. Donaldson. Aspects of the persistence of foot-and-mouth disease virus in animals–the carrier problem. *Microbes and Infection*, 4 (10):1099 – 1110, 2002a.

S. Alexandersen, Z. Zhang, S. M. Reid, G. H. Hutchings, and A. I. Donaldson. Quantities of infectious virus and viral rna recovered from sheep and cattle experimentally infected with foot-and-mouth disease virus o uk 2001. *Journal of General Virology*, 83(Pt 8):1915–1923, Aug 2002b.

S. Alexandersen, Z. Zhang, A. Donaldson, and A. Garland. The pathogenesis and diagnosis of foot-and-mouth disease. *Journal of Comparative Pathology*, 129(1):1–36, 2003.

I. Anderson. *Lessons to be Learned Inquiry.* House of commons, July 2002.

I. Anderson. *Foot and Mouth Disesease 2007: A review and lessons learned.* House of Commons, March 2008.

R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control (Oxford Science Publications).* Oxford University Press, USA, 1992. ISBN 019854040X.

Anon. Northumberland Report .I. *Veterinary Record*, 84(19):471–&, 1969.

Arizona-Software. Graphclick graph digitizer software., 2008. URL `www.arizona-software.ch`.

M. E. Arnold, D. J. Paton, E. Ryan, S. J. Cox, and J. W. Wilesmith. Modelling studies to estimate the prevalence of foot-and-mouth disease carriers after reactive vaccination. *Proceedings of Biological Sciences*, 2007. doi: 10.1098/rspb.2007.1154. URL `http://dx.doi.org/10.1098/rspb.2007.1154`.

L. Audige, M. G. Doherr, R. Hauser, and M. D. Salman. Stochastic modelling as a tool for planning animal-health surveys and interpreting screening-test results. *Preventive Veterinary Medicine*, 49(1-2):1–17, 2001.

AVIS Consortium. AVIS FMD 2002, 2002. URL `http://aleffgroup.com/avisfmd/`.

H. G. Bahnemann. Binary Ethylenimine as an Inactivant for Foot-and-Mouth-Disease Virus and Its Application for Vaccine Production. *Archives of Virology*, 47(1):47–56, 1975.

F. C. Baldock. What constitutes freedom from disease in livestock? *Australian Veterinary Journal*, 76(8):544–545, 1998.

B. M. Bankowski, R. Howey, N. Juleff, D. Gibson, C. Wright, S. J. Cox, P. V. Barnett, M. E. Woolhouse, and B. Charleston. Defining and quantifying the period of infectiousness in cattle experimentally infected with type O (UKG/2001) foot-and-mouth disease virus. 2008.

S. Baron, editor. *Medical Microbiology*. University of Texas Medical Branch, 1996. ISBN 0963117211.

A. D. S. Bastos, E. C. Anderson, R. G. Bengis, D. F. Keet, H. K. Winterbach, and G. R. Thomson. Molecular epidemiology of SAT3-type foot-and-mouth disease. *Virus Genes*, 27(3):283–290, 2003.

T. Bates, M. Thurmond, and T. Carpenter. Description of an epidemic simulation model for use in evaluating strategies to control an outbreak of foot-and-mouth disease. *American Journal of Veterinary Research*, 64(2):195–204, FEB 2003.

T. W. Bates, T. E. Carpenter, and M. C. Thurmond. Benefit-cost analysis of vaccination and preemptive slaughter as a means of eradicating foot-and-mouth disease. *American Journal of Veterinary Research*, 64(7):805–812, 2003.

BBC News Online. Farm crisis 'out of control', March 2001. URL `http://news.bbc.co.uk/1/hi/uk/1237623.stm`.

I. E. Bergmann, V. Malirat, L. E. Dias, and R. Dilandro. Identification of foot-and-mouth disease virus-free regions by use of a standardized enzyme-linked immuno-electrotransfer blot assay. *American Journal of Veterinary Research*, 57(7):972–974, 1996.

E. Blanco, L. J. Romero, M. El Harrach, and J.-M. Sanchez-Vizcaino. Serological evidence of fmd subclinical infection in sheep population during the 1999 epidemic in morocco. *Veterinary Microbiology*, 85(1):13–21, 2002.

A. Bouma, A. Dekker, and M. C. M. de Jong. No foot-and-mouth disease virus transmission between individually housed calves. *Veterinary Microbiology*, 98(1):29–36, 2004. ISSN 0378-1135 (Print).

A. J. Branscum, I. A. Gardner, and W. O. Johnson. Bayesian modeling of animal- and herd-level prevalences. *Preventive Veterinary Medicine*, 66(1-4):101–112, 2004.

A. J. Branscum, W. O. Johnson, and I. A. Gardner. Sample size calculations for disease freedom and prevalence estimation surveys. *Statistics in Medicine*, 25(15):2658–2674, Aug 2006.

E. Brocchi, I. E. Bergmann, A. Dekker, D. J. Paton, D. J. Sammin, M. Greiner, S. Grazioli, F. De Simone, H. Yadin, B. Haas, N. Bulut, V. Malirat, E. Neitzert, N. Goris, S. Parida, K. Sorensen, and K. De Clercq. Comparative evaluation of six elisas for the detection of antibodies to the non-structural proteins of foot-and-mouth disease virus. *Vaccine*, 24(47-48):6966–6979, November 2006.

B. M. D. Bronsvoort, A. D. Radford, V. N. Tanya, C. Nfon, R. P. Kitching, and K. L. Morgan. Molecular epidemiology of foot-and-mouth disease viruses in the Adamawa Province of Cameroon. *Journal of Clinical Microbiology*, 42(5):2186–2196, 2004.

B. M. d. C. Bronsvoort, J. Anderson, A. Corteyn, P. Hamblin, R. P. Kitching, C. Nfon, V. N. Tanya, and K. L. Morgan. Geographical and age-stratified distributions of foot-and-mouth disease virus-seropositive and probing-positive cattle herds in the adamawa province of cameroon. *Veterinary Record*, 159(10):299–+, SEP 2 2006.

S. P. Brooks and A. Gelman. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455, 1998. ISSN 10618600. URL `http://www.jstor.org/stable/1390675`.

F. Brown and J. Crick. Application of Agar-Gel Diffusion Analysis to a Study of the Antigenic Structure of Inactivated Vaccines Prepared from the Virus of Foot-and-Mouth Disease. *Journal of Immunology*, 82(5):444–447, 1959.

N. Bulut. Vaccination coverage data - SAP Institute Ankara, Turkey. (unpublished data), March 2006.

R. Burrows. The persistence of foot-and mouth disease virus in sheep. *Journal of Hygiene (Lond)*, 66(4):633–40, Dec 1968.

A. Cameron. Angus Cameron PhD Thesis, Chapter 4. Department of Geographical Sciences and Planning, Universty of Queensland, November 1997.

A. R. Cameron and F. C. Baldock. A new probability formula for surveys to substantiate freedom from disease. *Preventive Veterinary Medicine*, 34(1):1–17, 1998a.

A. R. Cameron and F. C. Baldock. Two-stage sampling in surveys to substantiate freedom from disease. *Preventive Veterinary Medicine*, 34(1):19–30, Feb. 1998b.

R. M. Cannon. Sense and sensitivity – designing surveys based on an imperfect test. *Preventive Veterinary Medicine*, 49(3-4):141–163, 2001.

R. M. Cannon. Demonstrating disease freedom–combining confidence levels. *Preventive Veterinary Medicine*, 52(3-4):227–249, 2002.

R. M. Cannon. Inspecting and monitoring on a restricted budget–where best to look? *Preventive Veterinary Medicine*, 92(1-2):163–74, Nov 2009. doi: 10.1016/j.prevetmed.2009.06.009.

R. M. Cannon and M. G. Garner. Assessing the risk of wind-borne spread of foot-and-mouth disease in australia. *Environment International*, 25(6-7):713 – 723, 1999.

P. B. Capstick, A. J. Garland, W. G. Chapman, and R. C. Masters. Production of Foot-and-Mouth Disease Virus Antigen from Bhk 21 Clone 13 Cells Grown and Infected in Deep Suspension Cultures. *Nature*, 205(4976):1135–&, 1965.

T. Carpenter, M. Thurmond, and T. Bates. A simulation model of intraherd transmission of foot and mouth disease with reference to disease spread before and after clinical diagnosis. *Journal of Veterinary Diagnostic Investigation*, 16(1):11–16, JAN 2004.

T. E. Carpenter. Epidemiologic modeling using a microcomputer spreadsheet package. *Am J Epidemiol*, 120(6):943–951, Dec 1984. ISSN 0002-9262 (Print).

T. E. Carpenter and I. A. Gardner. Simulation modeling to determine herd-level predictive values and sensitivity based on individual-animal test sensitivity and specificity and sample size. *Preventive Veterinary Medicine*, 27(1-2):57–66, 1996.

G. Celeux, F. Forbes, C. Robert, and D. Titterington. Deviance Information Criteria for Missing Data Models, 2006. URL `http://www.ceremade.dauphine.fr/~xian/cfrt05.pdf`.

Y. K. Choi, W. O. Johnson, and M. C. Thurmond. Diagnosis using predictive probabilities without cut-offs. *Statistics in Medicine*, 25(4):699–717, Feb. 2006.

E. M. Cottam, G. Thebaud, J. Wadsworth, J. Gloster, L. Mansley, D. J. Paton, D. P. King, and D. T. Haydon. Integrating genetic and epidemiological data to determine transmission pathways of foot-and-mouth disease virus. *Proceedings of Biological Sciences*, 275(1637):887–895, April 2008a.

E. M. Cottam, J. Wadsworth, A. E. Shaw, R. J. Rowlands, L. Goatley, S. Maan, N. S. Maan, P. P. C. Mertens, K. Ebert, Y. Li, E. D. Ryan, N. Juleff, N. P. Ferris, J. W. Wilesmith, D. T. Haydon, D. P. King, D. J. Paton, and N. J. Knowles. Transmission pathways of foot-and-mouth disease virus in the united kingdom in 2007. *PLoS Pathogens*, 4(4):e1000050, April 2008b.

M. Cowles and B. Carlin. Markov chain Monte Carlo convergence diagnostics: A comparative review, 1996.

M. J. Crawley. *The R Book*. Wiley, 2007. ISBN 0470510242.

P. S. Dawe, K. Sorensen, N. P. Ferris, I. T. Barnett, R. M. Armstrong, and N. J. Knowles. Experimental transmission of foot-and-mouth disease virus from carrier african buffalo (syncerus caffer) to cattle in zimbabwe. *Vet Rec*, 134(9):211–5, Feb 1994.

DEFRA. June Agricultural Survey, 2000. URL `http://www.defra.gov.uk/esg/work_htm/publications/cs/farmstats_web/Census/introduction.htm`.

DEFRA. Foot and mouth disease (serotypes A & O) in Turkey - VITT 1200/FMD - TURKEY. Technical report, DEFRA, December 2005. International Animal Health Division: Preliminary Outbreak Assessment.

DEFRA. Foot and mouth disease (Type A) in Turkey - VITT 1200/FMD – Turkey. Technical report, DEFRA, June 2006a. International Animal Health Division: Preliminary Outbreak Assessment.

DEFRA. Foot and mouth disease (type a) in turkey - vitt 1200/fmd – Turkey. Technical report, DEFRA, February 2006b. International Animal Health Division: Preliminary Outbreak Assessment.

DEFRA. Foot and mouth disease - Turkey - VITT 1200/FMD-Turkey. Technical report, DEFRA, January 2007a. International Animal Health Division: Preliminary Outbreak Assessment.

DEFRA. Foot and mouth disease – Turkey – an update - VITT 1200/FMD-Turkey. Technical report, DEFRA, March 2007b. International Animal Health Division: Preliminary Outbreak Assessment.

DEFRA. Foot and mouth disease – Turkey – Thrace region - VITT 1200/FMD-Turkey. Technical report, DEFRA, April 2007c. International Animal Health Division: Preliminary Outbreak Assessment.

DEFRA. Foot and mouth disease in the United Kingdom 2001 report to the OIE Part I: Great Britain. Technical report, 2002.

DEFRA. Notifiable disease archives, June 2008. URL `http://www.defra.gov.uk/animalh/diseases/control/archive.htm`.

DEFRA. History of disease control in the UK, 2009. URL `http://www.defra.gov.uk/foodfarm/farmanimal/diseases/control/history.htm`.

M. Denwood. *runjags: Run Bayesian MCMC Models in the BUGS syntax from Within R*, 2008. URL `http://cran.r-project.org/web/packages/runjags/`.

P. Desmettre. A century of veterinary vaccinology: the merieux initiative. *Archives of Virology*, 140(12):2293–2301, 12 1995.

P. J. Diggle. Spatio-temporal point processes, partial likelihood, foot and mouth disease. *Statistical Methods in Medical Research*, 15(4):325–336, Aug 2006.

T. R. Doel. FMD vaccines. *Virus Research*, 91(1):81–99, 2003.

A. W. Donald, I. A. Gardner, and A. D. Wiggins. Cutoff Points for Aggregate Herd Testing in the Presence of Disease Clustering and Correlation of Test Errors. *Preventive Veterinary Medicine*, 19(3-4):167–187, 1994.

B. Engel, W. Buist, K. Orsel, A. Dekker, K. de Clercq, S. Grazioli, and H. van Roermund. A bayesian evaluation of six diagnostic tests for foot-and-mouth disease for vaccinated and non-vaccinated cattle. *Preventive Veterinary Medicine*, 2008.

EUFMD. Proceedings of the 63rd session of the executive committee, Nov 1999. URL `http://www.fao.org/ag/againfo/commissions/en/eufmd/archive/execom_63report_txt.html#item3`.

EUFMD advisory group. Meetings of the advisory group on sero-surveillance post-outbreak for the thrace region. Technical report, EUFMD Research Group, June 2006. URL `http://www.fao.org/AG/AGAInfo/commissions/docs/excom73/App07.pdf`.

European Commission. General guidance on EU import and transit rules for live animals and animal products from third countries, 2006. URL `http://ec.europa.eu/food/international/trade/guide_thirdcountries2006_en.pdf`.

European Council. 2008/157/ec: Council decision of 18 February 2008 on the principles, priorities and conditions contained in the accession partnership with the republic of turkey and repealing decision 2006/35/ec, 2008. URL `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:051:0004:01:EN:HTML`.

N. M. Ferguson, C. A. Donnelly, and R. M. Anderson. The foot-and-mouth epidemic in Great Britain: Pattern of spread and impact of interventions. *Science*, 292(5519): 1155–1160, 2001.

J. Francis, J. Hawley, and T. Scott. Report on a survey of the effects of the 2001 Foot & Mouth Disease outbreak, (England), 2002. URL `http://statistics.defra.gov.uk/esg/reports/fmdreport.pdf`. ADAS Consulting Ltd, Wolverhampton.

I. A. Gardner, H. Stryhn, P. Lind, and M. T. Collins. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. *Preventive Veterinary Medicine*, 45(1-2):107–122, 2000.

C. J. Geyer. On the bogosity of MCMC diagnostics, 2005. URL `http://www.stat.umn.edu/~charlie/mcmc/diag.html`.

J. C. Gibbens and J. W. Wilesmith. Temporal and geographical distribution of cases of foot-and-mouth disease during the early weeks of the 2001 epidemic in great britain. *Veterinary Record*, 151(14):407–412, Oct 2002.

J. C. Gibbens, C. E. Sharpe, J. W. Wilesmith, L. M. Mansley, E. Michalopoulou, J. B. Ryan, and M. Hudson. Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months. *Veterinary Record*, 149(24):729–743, Dec 2001.

D. Gillespie. Exact stochastic simulation of coupled chemical-reactions. In *Abstracts of Papers of the American Chemical Society*, volume 173, pages 128–128, 1977.

J. Gloster, R. F. Sellers, and A. I. Donaldson. Long-Distance Transport of Foot-and-Mouth-Disease Virus over the Sea. *Veterinary Record*, 110(3):47–52, 1982.

H. Goldstein and P. McDonald, R. A general model for the analysis of multilevel data. *Psychometrika*, 53(4):455–467, 1988.

M. Greiner and A. Dekker. On the surveillance for animal diseases in small herds. *Preventive Veterinary Medicine*, 70(3-4):223–234, 2005.

M. J. Grubman and B. Baxt. Foot-and-mouth disease. *Clinical Microbiology Reviews*, 17(2):465–+, 2004.

T. E. Hanson, W. O. Johnson, I. A. Gardner, and M. P. Georgiadis. Determining the infection status of a herd. *Journal of Agricultural Biological and Environmental Statistics*, 8(4):469–485, 2003.

D. Harel. *Algorithmics: The Spirit of Computing*. Addison Wesley Publishing Company, 1987. ISBN 0201192403.

R. S. Hedger. Observations on the carrier state and related antibody titres during an outbreak of foot-and-mouth disease. *Journal of Hygiene (Lond)*, 68(1):53–60, Mar 1970.

P. Heidelberger and P. D. Welch. A spectral method for confidence interval generation and run length control in simulations. *Communications of the Association for Computing Machinery*, 24(4):233–245, 1981. ISSN 0001-0782. doi: http://doi.acm.org/10.1145/358598.358630.

A. Hindmarsh. Odepack, a systematized collection of ode solvers. In R. e. a. Stepleman, editor, *Scientific Computing: Applications of Mathematics and Computing to the Physical Sciences*, volume 1 of *IMACS Transactions on Scientific Computing*, pages 55–64, Amsterdam, Netherlands; New York, U.S.A., 1983. North-Holland.

HMSO. Report of the committee of inquiry on foot-and-mouth disease. Technical report, Her Majesty's Stationery Office, December 1969.

L. Hodgson-Jones. Interview. Easter Bush Veterinary Centre, Aug 2008.

M. Hohle and E. Jorgensen. Decision making based on sampled disease occurrence in animal herds. *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Proceeding*, 2711:220–229, 2003.

S. L. Hui and S. D. Walter. Estimating the error rates of diagnostic tests. *Biometrics*, 36(1):167–171, 1980.

R. W. Humphry, A. Cameron, and G. J. Gunn. A practical approach to calculate sample size for herd prevalence surveys. *Preventive Veterinary Medicine*, 65(3-4): 173–188, 2004.

Intervet. Foot and Mouth Disease vaccines manufactured by Intervet, 2005. URL `http://www.foot-and-mouth-disease.com/Product_information/intervet_vaccines.asp`.

Intervet. Checkit FMD - 3ABC Product Information, 2008. URL `http://www.foot-and-mouth-disease.com/Product_information/Checkit_FMD_test.asp`.

S. Jeffery. Foot and mouth 'killed people in 1800s'. The Guardian Online, November 2001. URL `http://www.guardian.co.uk/uk/2001/nov/23/footandmouth.simonjeffery`.

C. Jewell, T. Kypraios, R. Christley, and G. Roberts. A novel approach to real-time risk prediction for emerging infectious diseases: A case study in avian influenza h5n1. *Preventive Veterinary Medicine*, 91(1):19 – 28, 2009a. ISSN 0167-5877. Special Issue: GisVet 2007.

C. P. Jewell, M. J. Keeling, and G. O. Roberts. Predicting undetected infections during the 2007 foot-and-mouth disease outbreak. *J R Soc Interface*, 6(41):1145–51, Dec 2009b. doi: 10.1098/rsif.2008.0433.

W. O. Johnson, C.-L. Su, I. A. Gardner, and R. Christensen. Sample Size Calculations for Surveys to Substantiate Freedom of Populations from Infectious Agents. *Biometrics*, 60(1):165–171, 2004.

M. J. Keeling and P. Rohani. *Modeling Infectious Diseases in Humans and Animals.* Princeton University Press, 2007. ISBN 0691116172.

M. J. Keeling, M. E. Woolhouse, D. J. Shaw, L. Matthews, M. Chase-Topping, D. T. Haydon, S. J. Cornell, J. Kappey, J. Wilesmith, and B. T. Grenfell. Dynamics of the 2001 uk foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science*, 294(5543):813–817, Oct 2001.

M. J. Keeling, M. E. J. Woolhouse, R. M. May, G. Davies, and B. T. Grenfell. Modelling vaccination strategies against foot-and-mouth disease. *Nature*, 421(6919):136–142, 2003.

R. Kitching, N. Knowles, A. Samuel, and A. Donaldson. Development of foot-and-mouth disease virus strain characterisation - a review. *Tropical Animal Health and Production*, 21(3):153–166, 09 1989.

R. P. Kitching. Clinical variation in foot and mouth disease: cattle. *Rev Sci Tech*, 21 (3):499–504, Dec 2002.

R. P. Kitching and G. J. Hughes. Clinical variation in foot and mouth disease: sheep and goats. *Rev Sci Tech*, 21(3):505–512, Dec 2002.

R. P. Kitching, A. M. Hutber, and M. V. Thrusfield. A review of foot-and-mouth disease with special consideration for the clinical and epidemiological factors relevant to predictive modelling of the disease. *Vet J*, 169(2):197–209, Mar 2005. doi: 10.1016/j.tvjl.2004.06.001.

J. Klein, U. Parlak, F. Ozyoruk, and L. S. Christensen. The molecular epidemiology of foot-and-mouth disease virus serotypes a and o from 1998 to 2004 in turkey. *BMC Veterinary Research*, 2:35, 2006.

D. Klinkenberg, J. de Bree, H. Laevens, and M. C. M. de Jong. Within- and between-pen transmission of classical swine fever virus: a new method to estimate the basic reproduction ratio from transmission experiments. *Epidemiology and Infection*, 128 (2):293–299, Apr 2002.

P. M. Lee. *Bayesian Statistics: An Introduction.* A Hodder Arnold Publication, 1997. ISBN 0340677856.

R. M. Lequin. Enzyme immunoassay (eia)/enzyme-linked immunosorbent assay (elisa). *Clinical Chemistry*, 51(12):2415–2418, December 2005. doi: 10.1373/clinchem.2005.051532.

B. W. J. Mahy. Introduction and history of foot-and-mouth disease virus. In *Foot and Mouth Disease Virus*, volume 288 of *Current Topics in Microbiology and Immunology*, pages 1–8. 2005.

F. O. Mardones, A. M. Perez, J. Sanchez, A. AlKhamis, and T. E. Carpenter. Disease stages of foot-and-mouth disease virus (FMDV) serotype O infection in cattle, swine and small ruminants. Poster — Epidemics Conference, Asilomar, California, USA, December 2008.

P. A. J. Martin, A. R. Cameron, K. Barfod, E. S. G. Sergeant, and M. Greiner. Demonstrating freedom from disease using multiple complex data sources 2: case study–classical swine fever in Denmark. *Preventive Veterinary Medicine*, 79(2-4):98–115, May 2007a.

P. A. J. Martin, A. R. Cameron, and M. Greiner. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. *Preventive Veterinary Medicine*, 79(2-4):71–97, May 2007b.

S. W. Martin, M. Shoukri, and M. A. Thorburn. Evaluating the Health-Status of Herds Based on Tests Applied to Individuals. *Preventive Veterinary Medicine*, 14 (1-2):33–43, 1992.

B. Martinez-Lupez, A. M. Perez, A. De la Torre, and J. M. Sanchez-Vizcaino-Rodriguez. Quantitative risk assessment of foot-and-mouth disease introduction into spain via importation of live animals. *Preventive Veterinary Medicine*, 86(1-2):43–56, 8 2008.

M. McLaws, C. Ribble, W. Martin, and C. Stephen. Factors associated with the clinical diagnosis of foot and mouth disease during the 2001 epidemic in the uk. *Preventive Veterinary Medicine*, 77(1-2):65–81, Nov 2006.

M. McLaws, C. Ribble, C. Stephen, B. McNab, and P. R. Barrios. Reporting of suspect cases of foot-and-mouth-disease during the 2001 epidemic in the uk, and the herd

sensitivity and herd specificity of clinical diagnosis. *Preventive Veterinary Medicine*, 78(1):12–23, Jan 2007.

P. Moonen, L. Jacobs, A. Crienen, and A. Dekker. Detection of carriers of foot-and-mouth disease virus among vaccinated cattle. *Veterinary Microbiology*, 103(3-4):151 – 160, 2004.

MRC Cambridge. DIC: Deviance Information Criterion - WinBUGS online help. URL `http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/dicpage.shtml#q12`.

M. Murphy, R. Meyer, C. Mebus, A. Schudel, and M. Rodriguez. Analysis of sites of foot-and-mouth-disease virus persistence in carrier cattle via the polymerase chain-reaction. *Archives of Virology*, 136(3-4):299–307, 1994.

National Audit Office. The 2001 Outbreak of Foot and Mouth Disease. Website. Accessed 16/12/2008, June 2002. URL `http://www.nao.org.uk/system_pages/idoc.ashx?docid=91890c6b-8b85-49f7-afdc-42f4b3f5a030&version=-1`.

National Health Service. Can people catch foot and mouth disease? 'Choices' patient information website. Accessed 16/12/2008., 2008. URL `http://www.nhs.uk/chq/Pages/2348.aspx?CategoryID=95&SubCategoryID=95`.

R. Nsubuga, I. McKendrick, J. Hughes, Garteth, and M. Woolhouse. Estimation of unobservable infection process characteristics of expermiental FMD virus infection in groups of sheep. 2008. (in preparation).

OIE. General Definitions, 2005a. URL `http://www.oie.int/eng/normes/MCODE/en%5Fchapitre%5F1.1.1.htm`.

OIE. Foot and Mouth Disease. Differential diagnosis. Website. Accessed 16/12/2008, April 2002. URL `http://www.oie.int/eng/maladies/fiches/a_A010.HTM`.

OIE. Guidelines for the Establishment or the regaining of recognition for a foot an mouth disease fee country or zone, 2005b.

OIE. *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*, 2009a. Chapter 2.1.5 Foot and Mouth disease.

OIE. *Terrestrial Animal Health Code*, 2009b. Chapter 8.5 Foot and Mouth disease.

OIE-WAHID. World Animal Health Information Database, Disease distribution map for FMD (January—June 2008), 2008. URL `http://www.oie.int/wahis/public.php?page=disease_status_map`.

K. Orsel, A. Dekker, A. Bouma, J. A. Stegeman, and M. C. M. de Jong. Vaccination against foot and mouth disease reduces virus transmission in groups of calves. *Vaccine*, 23(41):4887–4894, 2005.

K. Orsel, M. C. M. de Jong, A. Bouma, J. A. Stegeman, and A. Dekker. The effect of vaccination on foot and mouth disease virus transmission among dairy cows. *Vaccine*, 25(2):327–335, 2007a.

K. Orsel, A. Dekker, A. Bouma, J. A. Stegeman, and M. C. M. de Jong. Quantification of foot and mouth disease virus excretion and transmission within groups of lambs with and without vaccination. *Vaccine*, 25(14):2673–2679, 2007b.

G. A. Paiba, J. Anderson, D. J. Paton, A. W. Soldan, S. Alexandersen, M. Corteyn, G. Wilsden, P. Hamblin, D. K. J. MacKay, and A. I. Donaldson. Validation of a foot-and-mouth disease antibody screening solid-phase competition elisa (spce). *Journal of Virological Methods*, 115(2):145–158, Feb 2004.

J. Parker. Presence and inactivation of foot-and-mouth disease virus in animal faeces. *Veterinary Record*, 88(25):659–&, 1971.

D. J. Paton, K. de Clercq, and A. Dekker. Post-Vaccinal serosurveillance for FMD: a European perspective on progres and problems. Website, October 2004. URL `http://www.fao.org/ag/againfo/commissions/en/eufmd/crete.html`.

D. J. Paton, K. de Clercq, M. Greiner, A. Dekker, E. Brocchi, I. Bergmann, D. J. Sammin, S. Gubbins, and S. Parida. Application of non-structural protein antibody tests in substantiating freedom from foot-and-mouth disease virus infection after emergency vaccination of cattle. *Vaccine*, 24(42-43):6503–6512, Oct 2006. doi: 10.1016/j.vaccine.2006.06.032.

M. Plummer. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing; Vienna, Austria., 2003.

M. Plummer, N. Best, K. Cowles, and K. Vines. CODA: Convergence diagnosis and output analysis for MCMC. In *R News*, volume 6 of *Rnews2006-1*, pages 7–11. Rnews2006-1, March 2006.

H. Prempeh, R. Smith, and B. Muller. Foot and mouth disease: the human consequences - the health consequences are slight, the economic ones huge. *British Medical Journal*, 322(7286):565–566, MAR 10 2001.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2007. URL `http://www.R-project.org`.

A. E. Raftery and S. Lewis. How many iterations in the gibbs sampler. In *In Bayesian Statistics 4*, pages 763–773. Oxford University Press, 1992.

K. Rainwater-Lovett, J. Pacheco, C. Packer, and L. Rodriguez. Detection of foot-and-mouth disease virus infected cattle using infrared thermography. *Veterinary Journal*, Feb 2008.

E. Ryan, J. Gloster, S. M. Reid, Y. Li, N. P. Ferris, R. Waters, N. Juleff, B. Charleston, B. Bankowski, S. Gubbins, J. W. Wilesmith, D. P. King, and D. J. Paton. Clinical and laboratory investigations of the outbreaks of foot-and-mouth disease in southern england in 2007. *Veterinary Record*, 163(5):139–147, AUG 2 2008.

M. Salmon. *Animal Disease Surveillance and Survey Systems: Methods and Applications*. Wiley-Blackwell, 2003. ISBN 0813810310.

P. Scott. Interview. Easter Bush Veterinary Centre, Midlothian, UK, Aug 2008.

A. E. Segarra and J. M. Rawson. Foot and mouth disease: A threat to u.s. agriculture, 2001. URL `http://www.nationalaglawcenter.org/assets/crs/RS20890.pdf`.

F. Sobrino. *Foot and Mouth Disease*. CRC Press, Boca Raton, 2004. ISBN 9780849329517.

J. H. Sorensen, D. K. Mackay, C. O. Jensen, and A. I. Donaldson. An integrated model to predict the atmospheric spread of foot-and-mouth disease virus. *Epidemiol Infect*, 124(3):577–590, Jun 2000.

K. J. Sorensen, K. G. Madsen, E. S. Madsen, J. S. Salt, J. Nqindi, and D. K. J. Mackay. Differentiation of infection from vaccination in foot-and-mouth disease by

the detection of antibodies to the non-structural proteins 3D, 3AB and 3ABC in ELISA using antigens expressed in baculovirus. *Archives of Virology*, 143(8):1461–1476, 1998.

K. J. Sorensen, K. de Stricker, K. C. Dyrting, S. Grazioli, and B. Haas. Differentiation of foot-and-mouth disease virus infected animals from vaccinated animals using a blocking ELISA based on baculovirus expressed FMDV 3ABC antigen and a 3ABC monoclonal antibody. *Archives of Virology*, 150(4):805–814, 2005.

D. J. Spiegelhalter, A. Thomas, and N. G. Best. *WinBUGS Version 1.2 User Manual*. MRC Biostatistics Unit, 1999.

D. J. Spiegelhalter, N. G. Best, B. R. Carlin, and A. van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 64:583–616, 2002.

M. Stephens. Dealing with label switching in mixture models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 62(4):795–809, 2000. ISSN 13697412.

G. Streftaris and G. J. Gibson. Bayesian analysis of experimental epidemics of foot-and-mouth disease. *Proceedings of Biological Sciences*, 271(1544):1111–1117, Jun 2004. doi: 10.1098/rspb.2004.2715.

E. A. Suess, I. A. Gardner, and W. O. Johnson. Hierarchical Bayesian model for prevalence inferences and determination of a country's status for an animal pathogen. *Preventive Veterinary Medicine*, 55(3):155–171, 2002.

P. Sutmoller and A. Gaggero. Foot-and-mouth disease carriers. *The Veterinary Record*, 77:968–969, 1965.

P. Sutmoller, J. W. McVicar, and G. E. Cottral. The epizootiological importance of foot-and-mouth disease carriers. i. experimentally produced foot-and-mouth disease carriers in susceptible and immune cattle. *Arch Gesamte Virusforsch*, 23(3):227–35, 1968.

TAIEX. *Workshop on the Design and Interpretation of Post FMD - Vaccination Sero-surveillance by NSP Tests*, 2007. Technical Assistance Information Exchange Instrument of the European Commision. Held at Veterinary and Agrochemical Research Centre, Brussels, 5-7 June 2007.

P. Thomson. Absence of specified animal diseases from New Zealand. Technical report, Biosecurity New Zealand - Ministry of Agriculture and Forestry - Nea Zealand, March 2009. URL `http://www.biosecurity.govt.nz/files/pests/surv-mgmt/surv/absence-specified-animal-diseases.pdf`.

M. Thrusfield, L. Mansley, P. Dunlop, A. Pawson, and J. Taylor. The foot-and-mouth disease epidemic in Dumfries and Galloway, 2001. 2: Serosurveillance, and efficiency and effectiveness of control procedures after the national ban on animal movements. *Veterinary Record*, 156(9):269–278, Feb 2005a.

M. Thrusfield, L. Mansley, P. Dunlop, J. Taylor, A. Pawson, and L. Stringer. The foot-and-mouth disease epidemic in Dumfries and Galloway, 2001. 1: Characteristics and control. *Veterinary Record*, 156(8):229–252, Feb 2005b.

M. C. Thurmond and A. M. Perez. Modeled detection time for surveillance for foot-and-mouth disease virus in bulk tank milk. *American Journal of Veterinary Research*, 67 (12):2017–2024, Dec 2006.

M. J. Tildesley, N. J. Savill, D. J. Shaw, R. Deardon, S. P. Brooks, M. E. J. Woolhouse, B. T. Grenfell, and M. J. Keeling. Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. *Nature*, 440(7080):83–86, Mar. 2006.

T. Tsutsui, N. Minami, M. Koiwai, T. Hamaoka, I. Yamane, and K. Shimura. A stochastic-modeling evaluation of the foot-and-mouth-disease survey conducted after the outbreak in Miyazaki, Japan in 2000. *Preventive Veterinary Medicine*, 61(1): 45–58, 2003.

A. G. J. Velthuis, A. Bouma, W. E. A. Katsma, G. Nodelijk, and M. C. M. De Jong. Design and analysis of small-scale transmission experiments with animals. *Epidemiology and Infection*, 135(2):202–217, 2007.

VPH Test standardisation group, DEFRA. Foot and mouth disease: Detection of antibodies against non-structural proteins. Technical report, NRM number 013, 2007. URL `http://www.defra.gov.uk/animalh/diseases/vetsurveillance/pdf/nrm013-fmd-protein.pdf`.

Wen-Bin Chung, Karl Johan Sorensen, Pei-Chih Liao, Pin-Cheng Yang, and Ming-Hwa Jong. Differentiation of foot-and-mouth disease virus-infected from vaccinated pigs by enzyme-linked immunosorbent assay using nonstructural protein 3AB as the

antigen and application to an eradication program. *Journal of Clinical Microbiology*, 40(8):2843–2848, Aug 2002.

M. S. Williams, E. D. Ebel, and S. J. Wells. Poisson sampling: A sampling strategy for concurrently establishing freedom from disease and estimating population characteristics. *Preventive Veterinary Medicine*, 89(1-2):34 – 42, 2009a.

M. S. Williams, E. D. Ebel, and S. J. Wells. Population inferences from targeted sampling with uncertain epidemiologic information. *Preventive Veterinary Medicine*, 89(1-2):25 – 33, 2009b.

R. Woodrow Setzer. *odesolve: Solvers for Ordinary Differential Equations*, 2008. R package: version 0.5-20.

A. Woods. *A Manufactured Plague? The History of Foot-and-mouth Disease in Britain.* Earthscan Publications, London, 2004. ISBN 9781844070800.

M. E. Woolhouse, D. T. Haydon, A. Pearson, and R. P. Kitching. Failure of vaccination to prevent outbreaks of foot-and-mouth disease. *Epidemiology and Infection*, 116(3): 363–371, June 1996.

World Trade Organisation. Understanding the wto agreement on sanitary and phytosanitary measures, 2008. URL `http://www.wto.org/english/tratop_e/sps_e/spsund_e.htm`.

Z. Zhang and S. Alexandersen. Detection of carrier cattle and sheep persistently infected with foot-and-mouth disease virus by a rapid real-time RT-PCR assay. *Journal of Virological Methods*, 111(2):95–100, Aug 2003.